



# Etude fonctionnelle de deux marqueurs régionaux du cerveau chez la souris

Nada Lepejova Caudy

## ► To cite this version:

Nada Lepejova Caudy. Etude fonctionnelle de deux marqueurs régionaux du cerveau chez la souris. Biologie animale. Université Paris Sud - Paris XI, 2011. Français. NNT : 2011PA112127 . tel-00868783

**HAL Id: tel-00868783**

**<https://theses.hal.science/tel-00868783>**

Submitted on 2 Oct 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**UNIVERSITE PARIS-SUD 11**

**ECOLE DOCTORALE : « GENES, GENOMES, CELLULES »**  
**Laboratoire de Physiogénomique**

*DISCIPLINE : BIOLOGIE ANIMALE*

**THÈSE DE DOCTORAT**  
soutenue le 09/09/2011

par

**Nad'a CAUDY née LEPEJOVÁ**

pour obtenir le grade de : **Docteur de l'Université Paris Sud-XI**

|   |
|---|
| <p><b>Etude fonctionnelle de deux marqueurs régionaux<br/>du cerveau chez la souris</b></p> |
|---|

**Composition du jury :**

|                             |                   |  |
|-----------------------------|-------------------|--|
| <i>Président du jury :</i>  | Pierre CAPY       | Professeur (Université Paris Sud-XI)   |
| <i>Rapporteurs :</i>        | Olivier LAPREVOTE | Professeur (Université Paris Descartes - Paris V)                            |
|                             | Laurent JOURNOT   | Directeur de recherche (Institut de Génétique<br>Fonctionnelle, Montpellier) |
| <i>Examineur :</i>          | François TRONCHE  | Directeur de recherche CNRS<br>(Université Pierre et Marie Curie - Paris VI) |
| <i>Directeur de thèse :</i> | Jean-Marc ELALOUF | Chercheur (CEA, Saclay)  |
| <i>Encadrant :</i>          | Michel de CHALDEE | Chercheur (CEA, Saclay)  |



« En vérité, le chemin importe peu, la volonté d'arriver suffit à tout. »

Albert Camus  
Extrait de *Le Mythe de Sisyphe*



Túto prácu venujem mojim rodičom,  
mojej najlepšej kamarátke Katke  
et  
à mon mari merveilleux.

## Remerciements

*Je remercie mon directeur de thèse Jean-Marc Elalouf de m'avoir accueilli au sein de son équipe et d'avoir mis à ma disposition tous les moyens techniques que possède son laboratoire.*

*Je tiens à remercier tout particulièrement Michel de Chaldée pour son encadrement rigoureux, pour les nombreuses discussions et ses nombreux conseils ainsi que pour sa disponibilité pour les diverses expériences.*

*Je remercie Christophe Junot pour la formidable collaboration qu'il nous a offerte durant mes trois années de thèse, pour sa gentillesse, son accueil toujours égal et chaleureux et le temps qu'il a consacré à ce projet malgré ses nombreuses responsabilités. Merci également à Geoffrey Madalinski pour son aide et assistance technique.*

*Je souhaite remercier l'équipe de Nicole Déglon et d'Emmanuel Brouillet pour le travail de qualité qu'ils ont accompli au cours de notre collaboration. Merci pour les conseils et les discussions fructueuses.*

*Je remercie Messieurs Olivier Laprêvotte et Laurent Journot d'avoir accepté de juger mon travail en tant que rapporteurs ainsi que les autres membres de mon jury de thèse, Pierre Capy et François Tronche, de m'avoir fait l'honneur de participer à ma soutenance.*

*Un grand merci à toutes les personnes de l'animalerie pour s'être occupés de nos souris transgéniques.*

*Je voudrais exprimer mes remerciements à Marie-Alix Poul pour ses conseils et son aide dans l'utilisation de la technique du phage display dans la production des anticorps, ainsi qu'à Françoise Ochsenbeim pour le temps qu'elle a consacré à nous aider à leur purification.*

*Je n'oublie pas de remercier Nicole Faucon-Biguet pour son aide et sa gentillesse de nous avoir appris à disséquer le ganglion cervical supérieur.*

*Merci aux filles de mon équipe Céline Bon et Véronique Berthonaud, pour la bonne ambiance dans le bureau, pour les discussions intéressantes et les conseils enrichissants ainsi que pour les encouragements dans les moments difficiles... Merci Céline pour cette odeur sublime de café tout les matins en arrivant !*

*Un merci tout spécial à Catherine Doreau pour sa toujours bonne humeur, sa gentillesse et son art de « secret, taire ». Tu me manqueras beaucoup !*

*Merci également à toutes les personnes du service qui ont rendue mon séjour dans cet établissement très agréable notamment Emmanuel Gabette pour m'avoir toujours fait rire, Gilles Lagniel pour les pauses café reposantes, Arounie Tavenet pour son amitié dans les moments meilleurs et pires et tous les autres...*

*Je voudrais remercier mes meilleurs amis du monde Katka, Jacquo, Mathieu, Camille et Julie qui ont toujours été à mes côtés pour me soutenir et pour partager tous les moments importants de ma vie...*

*Ďakujem mojej rodine, za ich lásku, dôveru a podporu. Ďakujem mojim rodičom, že ma nikdy nenechali vzdať sa, bez vás by som to nikdy tak d'aleko nedotiahla!*

*Un énorme merci à mon mari Nicolas, pour le soutien, et pour avoir supportée mes humeurs tout au long de ma thèse et surtout pendant la rédaction de ce manuscrit, pour ne pas m'avoir laissé plonger dans mon désespoir... Merci également pour ton aide graphique et la relecture fréquente, pour le français qui n'est toujours pas une langue évidente... Merci pour tout et pour être toujours là pour moi.*





# TABLE DES MATIERES

|   |           |
|---|-----------|
| <b>FIGURES.....</b>   | <b>4</b>  |
| <b>TABLES.....</b>  | <b>5</b>  |
| <b>ABREVIATIONS .....</b>   | <b>6</b>  |
| <b>I. Contexte au sein du laboratoire.....</b>  | <b>12</b> |
| <b>II. Introduction générale.....</b>   | <b>13</b> |
| <b>1. Biomarqueurs .....</b>  | <b>13</b> |
| 1.1. Différentes approches technologiques pour la découverte de nouveaux biomarqueurs moléculaires..... | 15        |
| 1.1.1. L'approche génomique.....  | 16        |
| 1.1.2. L'approche épigénomique.....   | 18        |
| 1.1.3. L'approche transcriptomique .....  | 21        |
| 1.1.4. Approche protéomique .....   | 24        |
| 1.1.5. Approche métabolomique .....   | 25        |
| 1.1.5.1. Profilage .....  | 25        |
| 1.1.5.1.1. Ionisation .....   | 26        |
| 1.1.5.1.2. Chromatographie .....  | 27        |
| 1.1.5.1.3. Analyseurs de masse .....  | 28        |
| 1.1.5.1.4. Analyse des données .....  | 29        |
| 1.1.5.2 Spectrométrie de masse en tandem .....  | 30        |
| 1.1.5.3 Applications .....  | 31        |
| <b>2. Caractérisation fonctionnelle des biomarqueurs .....</b>  | <b>32</b> |
| 2.1. Invalidation/surexpression.....  | 32        |
| 2.1.1. Invalidation génique chez la souris .....  | 32        |
| 2.1.1.1. La souris comme organisme modèle .....   | 32        |
| 2.1.1.1.1. Histoire de l'utilisation de la souris en génétique .....                                    | 32        |
| 2.1.1.1.2. Intérêt de ce modèle pour la recherche.....  | 33        |
| 2.1.1.2. Production de souris génétiquement modifiées.....  | 33        |
| 2.1.1.3. Méthodes d'inactivation génique chez la souris .....   | 35        |
| 2.1.1.3.1. Knock-out et knock-in.....   | 35        |
| 2.1.1.3.2. Knock-out conditionnel.....  | 37        |
| 2.1.1.3.2.1. Contrôle dans l'espace.....  | 37        |
| 2.1.1.3.2.2. Contrôle dans le temps .....   | 37        |
| 2.1.1.3.3. Inactivation à grande échelle.....   | 40        |

|  |                |
|--|----------------|
| 2.1.1.3.3.1. Mutagénèse par l'ENU .....                                | 40             |
| 2.1.1.3.3.2. Approche Genetrap .....                                   | 41             |
| 2.1.1.4. Lignées congéniques .....                                     | 44             |
| 2.1.2. Knock-down .....  | 45             |
| 2.1.3. Surexpression .....   | 47             |
| 2.2. Analyse spatio-temporelle de l'expression d'un gène .....         | 48             |
| 2.2.1. Approches moléculaires et biochimiques.....                     | 48             |
| 2.2.1.1. A l'échelle de l'ARN.....                                     | 48             |
| 2.2.1.2. A l'échelle de la protéine .....                              | 51             |
| 2.2.2. Analyses histologiques .....                                    | 51             |
| 2.2.2.1. A l'échelle de l'ARN.....                                     | 51             |
| 2.2.2.2. A l'échelle de la protéine .....                              | 52             |
| 2.2.2.3. Marquage Xgal .....   | 52             |
| 2.3 Autres approches .....   | 53             |
| 2.3.1 Analyse bioinformatique .....                                    | 53             |
| 2.3.2. Recherche d'interactants .....                                  | 54             |
| 2.3.2.1. Double hybride .....  | 54             |
| 2.3.2.2. Autres techniques .....                                       | 57             |
| <b>3. Récapitulation.....</b>  | <b>58</b>      |
| <br><b>III. Etude fonctionnelle de la Capucine.....</b>                | <br><b>59</b>  |
| <br><b>1. Le striatum .....</b>  | <br><b>59</b>  |
| 1.1. Structure .....   | 59             |
| 1.2. Fonction.....   | 62             |
| 1.2.1. Fonction du striatum au sein des ganglions de la base.....      | 63             |
| 1.2.2. Les différents circuits striataux des ganglions de la base..... | 64             |
| <br><b>2. La maladie de Huntington.....</b>                            | <br><b>65</b>  |
| 2.1. Symptômes .....   | 65             |
| 2.2. Neuropathologie .....   | 67             |
| 2.3. Génétique.....  | 68             |
| 2.4. Physiopathologie .....  | 69             |
| 2.5. Traitement de la maladie .....                                    | 70             |
| 2.6. Modèles de la maladie de Huntington.....                          | 72             |
| 2.6.1. Souris transgéniques.....                                       | 72             |
| 2.6.2. Les modèles lentiviraux.....                                    | 73             |
| <b>3. Article N° 1.....</b>  | <b>74</b>      |
| <br><b>4. Production d'anticorps contre la Capucine.....</b>           | <br><b>101</b> |
| <br><b>5. Discussion .....</b>   | <br><b>106</b> |

|   |            |
|---|------------|
| <b>IV. Etude fonctionnelle d'AGPAT4.....</b>                                | <b>108</b> |
| <b>1. Métabolisme des glycérophospholipides .....</b>                       | <b>108</b> |
| 1.1. Fonction des glycérophospholipides .....                               | 109        |
| 1.2. Biosynthèse des glycérophospholipides.....                             | 111        |
| 1.2.1. La voie de Kennedy.....  | 111        |
| 1.2.2. Le Cycle de Land .....   | 111        |
| 1.3. Les membres de la famille Agpat.....                                   | 112        |
| 1.3.1. Fonctions des différents membres de la famille Agpat .....           | 114        |
| 1.3.1.1. Membres les mieux caractérisés.....                                | 115        |
| 1.3.1.2. Membres présentant les homologies les plus fortes avec Agpat4..... | 118        |
| 1.3.2. Profil d'expression des Agpats 1 à 5 et 8 .....                      | 119        |
| <b>2. Catécholamines .....</b>  | <b>120</b> |
| 2.1. Sites de synthèse et principales fonctions .....                       | 120        |
| 2.1.1. La dopamine .....  | 120        |
| 2.1.1.1. Système nigrostrié .....   | 121        |
| 2.1.1.2. Système mésolimbique.....  | 122        |
| 2.1.1.3. Système mésocortical .....   | 125        |
| 2.1.2. La noradrénaline et l'adrénaline.....                                | 125        |
| 2.2. Métabolisme des catécholamines .....                                   | 126        |
| 2.2.1. Biosynthèse .....  | 126        |
| 2.2.2. Dégradation .....  | 127        |
| <b>3. Article N°2.....</b>  | <b>129</b> |
| <b>4. Discussion .....</b>  | <b>164</b> |
| <b>BIBLIOGRAPHIE .....</b>  | <b>168</b> |

## FIGURES

|   |     |
|---|-----|
| <b>Figure 1</b> : Principe de base des GWAS (genome wide association studies).....  | 17  |
| <b>Figure 2</b> : Représentation schématique de la méthode ChIP-seq. ....   | 20  |
| <b>Figure 3</b> : Illustration de la méthode qui combine les résultats des GWAS et EWAS. ....                                     | 20  |
| <b>Figure 4</b> : Le principe de la technique des puces à ADN.....  | 22  |
| <b>Figure 5</b> : Schéma de la technique SAGE utilisée au laboratoire. ....   | 23  |
| <b>Figure 6</b> : Schéma général des étapes de fonctionnement d'un spectromètre de masse .....                                    | 26  |
| <b>Figure 7</b> : Schéma de la MS/MS. ....  | 31  |
| <b>Figure 8</b> : Représentation schématique de la production de souris génétiquement modifiées.<br>.....                         | 34  |
| <b>Figure 9</b> : Principe général de l'inactivation d'un gène par recombinaison homologue.....                                   | 36  |
| <b>Figure 10</b> : Mutagenèse conditionnelle basée sur le système Cre/loxP. ....  | 38  |
| <b>Figure 11</b> : Systèmes Tet-Off et Tet-On. ....   | 39  |
| <b>Figure 12</b> : Principe de la technique « gene trap ». ....   | 42  |
| <b>Figure 13</b> : Un vecteur de « gene trap » amélioré. ....   | 43  |
| <b>Figure 14</b> : Processus de « congénisation » des souris transgéniques.....   | 44  |
| <b>Figure 15</b> : Vue schématique du knock-down par l'ARN interférence. ....   | 46  |
| <b>Figure 16</b> : Différences entre les techniques TaqMan et SYBR-Green. ....  | 50  |
| <b>Figure 17</b> : Système du double hybride.....   | 56  |
| <b>Figure 18</b> : Etude intégrée de la fonction d'un gène. ....  | 58  |
| <b>Figure 19</b> : Schéma de la régulation de la phosphorylation de DARPP-32 par la dopamine. ....                                | 61  |
| <b>Figure 20</b> : Schéma du circuit cortico-striato-thalamo-cortical. ....   | 64  |
| <b>Figure 21</b> : Fingerprint des gènes de scFv. ....  | 102 |
| <b>Figure 22</b> : Séquences des scFv correspondant aux différents profils obtenus par<br>fingerprinting.....                     | 103 |
| <b>Figure 23</b> : Test d'affinité des scFv pour un fragment recombinant de Capucine par ELISA.<br>.....                          | 104 |
| <b>Figure 24</b> : Production d'un scFv du profil « A ». ....   | 105 |
| <b>Figure 25</b> : Structure des glycérophospholipides. ....  | 109 |
| <b>Figure 26</b> : Les voies de biosynthèse des glycérophospholipide .....  | 112 |
| <b>Figure 27</b> : Les fonctions de différents membres de la famille AGPAT dans la biosynthèse<br>des glycérophospholipides ..... | 115 |
| <b>Figure 28</b> : Arbre phylogénétique de différents membres de la famille AGPAT chez la souris<br>.....                         | 118 |
| <b>Figure 29</b> : La synapse dopaminergique .....  | 123 |
| <b>Figure 30</b> : Biosynthèse des catécholamines.....  | 127 |
| <b>Figure 31</b> : Dégradation des catécholamines (D'après KEGG) .....  | 128 |

## TABLES

|   |     |
|---|-----|
| <b>Tableau 1:</b> Les principales chaînes d'acides gras .....                       | 109 |
| <b>Tableau 2 :</b> Les motifs conservés dans les différentes acyltransférases ..... | 113 |

## ABBREVIATIONS

|               |  |
|---------------|--|
| 3-MT :        | 3-méthoxytyramine  |
| Acyl-coA :    | Acyl-coenzyme A  |
| ADN :         | Acide désoxyribonucléique  |
| ADNc :        | Acide désoxyribonucléique complémentaire                               |
| AE :          | Enzyme d'ancrage   |
| Agpat:        | 1-acylglycerol-3-phosphate O-acyltransferase                           |
| ALCAT :       | Acyl-CoA:lysocardiolipin acyltransférase                               |
| AMPA:         | $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid           |
| AMPc :        | Adénosine mono-phosphate cyclique                                      |
| Ar :          | Argon  |
| ARNi:         | Acide ribonucléique interférent  |
| ARNm :        | Acide ribonucléique messenger  |
| ATG :         | Adénine thymine guanine  |
| BDNF :        | Brain-Derived Neurotrophic Factor                                      |
| BGEM :        | Brain Gene Expression Map  |
| $\beta$ geo : | Galactosidase/neomycin phosphotransferase fusion gene                  |
| BLAST :       | Basic Local Alignment Search Tool                                      |
| Btk :         | Burton's tyrosine kinase   |
| C-terminal :  | Terminaison carbone  |
| CAG :         | Cytosine adénine guanine   |
| CalDAG-GEFI : | Calcium-Diacylglycerol Guanine Nucleotide Exchange Factor I            |
| CDP-DAG :     | Cytidine diphospho diacylglycérol                                      |
| CGL :         | Lipodystrophie généralisée congénitale                                 |
| ChIP :        | Chromatin immunoprecipitation  |
| ChIP-seq :    | Chromatin immunoprecipitation coupled to massively parallel sequencing |
| CHO :         | Chinese Hamster Ovary  |
| CI :          | Ionisation chimique  |
| CID :         | Collision-induced dissociation   |
| CL :          | Cardiolipine   |
| CMV :         | Cytomegalovirus immediate-early promoter                               |
| coIP :        | Co-immunoprécipitation   |

|            |   |
|------------|---|
| COMT :     | Catéchol-O-méthyltransférase  |
| COS-1 :    | Cells being CV-1 (simian) in Origin, and carrying the SV40 genetic material |
| CpG :      | Cytosine phosphodiester guanine   |
| Cre :      | Causes recombinaison  |
| Crym :     | Crystalline mu  |
| cTnI :     | Troponine cardiaque I   |
| DARPP-32 : | Dopamine- and cAMP-Regulated PhosphoProtein, 32kDa                          |
| DAT :      | Dopamine active transporter   |
| DBA :      | Dilute Brown non-Agouti   |
| DHAPAT :   | Dihydroxyacetonephosphate acyltransferase                                   |
| DIP :      | Database of Interacting Proteins  |
| Dnmt :     | DNA (cytosine-5-)-methyltransferase 1                                       |
| DOPA :     | 3,4-dihydroxyphénylalanine  |
| DOPAC :    | Acide 3,4-dihydroxyphenylacétique   |
| DOPAL :    | 3,4-dihydroxyphénylacétaldéhyde   |
| dT :       | Désoxythymidine   |
| EBI :      | European Bioinformatics Institute   |
| EF1A :     | Human elongation factor 1 $\alpha$ promoter                                 |
| EGFP :     | Enhanced Green Fluorescent Protein  |
| EI :       | Ionisation électronique   |
| ELISA :    | Enzyme-linked immunosorbent assay   |
| EMBL :     | European Molecular Biology Laboratory                                       |
| ENU :      | N-éthyl-N-nitrosourée   |
| ES :       | Embryonic stem  |
| ESI :      | Electrospray ionization   |
| ESI-MS :   | Electrospray ionisation-mass spectrometry                                   |
| EST :      | Expressed sequence tag  |
| EWAS :     | Epigenome-wide association studies  |
| EYFP :     | Enhanced Yellow Fluorescent Protein   |
| FAAH :     | Fatty acid amide hydrolase  |
| FAB :      | Bombardement par atomes rapides   |
| FACS :     | Fluorescence-Activated Cell Sorting   |
| FBN2 :     | Fibrillin-2   |



|                |   |
|----------------|---|
| Flp :          | Flippase recombination enzyme   |
| FokI :         | Flavobacterium okeanokoites I   |
| FRT :          | Flippase Recognition Target   |
| FTO :          | Fat mass and obesity-associated gene  |
| GABA :         | Acide gamma-aminobutyrique  |
| GC :           | Chromatographie en phase gazeuse  |
| GC-MS :        | Gaz chromatography-mass spectrometry  |
| Genbank :      | Gene Bank   |
| GENSAT :       | Gene Expression Nervous System Atlas  |
| GFP :          | Green Fluorescent Protein   |
| GPAT :         | Glycerol-3-phosphate acyltransferase  |
| GWAS :         | Genome wide association studies   |
| HA :           | Hémagglutinine  |
| HapMap :       | Haplotype map   |
| HCMV :         | CytomégaloVirus humain  |
| HDAC :         | Histone désacétylase  |
| He :           | Hélium  |
| HeLa :         | Cellules cancéreuses prélevées sur Henrietta Lacks atteinte d'un cancer de l'utérus |
| HMDB :         | The Human Metabolome Database   |
| Htt :          | Huntingtine   |
| HVA :          | Acide homovanillique  |
| IMGT/V-QUEST : | The international ImMunoGeneTics information system/ V-QUERy and Standardization    |
| IMPC :         | International Mouse Phenotyping Consortium  |
| IKMC :         | International Knockout Mouse Consortium   |
| JNK :          | c-Jun N-terminal Kinase   |
| kDa :          | kilodalton  |
| KEGG :         | Kyoto Encyclopedia of Genes and Genomes   |
| Kr :           | Krypton   |
| LacZ :         | Lactose Z (codant pour la $\beta$ -Galactosidase)                                   |
| LC-MS :        | Liquid chromatography-mass spectrometry   |
| LCLAT :        | Lysocardiolipine acyltransférase  |
| L-DOPA :       | L-dihydroxyphénylalanine / Levodopa   |

|                |  |
|----------------|--|
| LoxP :         | locus of X-over P1   |
| LPA :          | Acide lysophosphatidique   |
| LPAAT :        | Lysophosphatidic acid acyltransferase  |
| LPC :          | Lysophosphatidylcholine  |
| LPCAT :        | Lysophosphatidylcholine acyltransferase                                      |
| LPE :          | Lysophosphatidylethanolamine   |
| LPEAT :        | Lysophosphatidylethanolamine acyltransferase;                                |
| LPG :          | Lysophosphatidylglycérol   |
| LPGAT :        | Lysophosphatidylglycérol acyltransférase                                     |
| LPI :          | Lysophosphatidylinositol   |
| LPIAT :        | Lysophosphatidylinositol acyltransférase                                     |
| LPL :          | Lysophospholipide  |
| LPLAT :        | Lysophospholipide acyltransférase  |
| LPS :          | Lysophosphatidylsérine   |
| LPSAT :        | Lysophosphatidylsérine acyltransférase                                       |
| LTQ :          | Linear quadrupole ion trap   |
| LysoPAFAT :    | Lyso-platelet-activating factor acetyltransferase                            |
| m/z :          | Rapport masse/charge   |
| MAPK :         | Mitogen-activated protein kinase   |
| MALDI :        | Matrix-assisted laser desorption/ionization                                  |
| MALDI-TOF-MS : | Matrix-Assisted Laser Desorption Ionization-Time Of Flight-Mass Spectrometry |
| MAO :          | Monoamine oxydase  |
| MBOAT :        | Membrane bound O-acyl transferase  |
| METLIN :       | Metabolite and Tandem MS Database  |
| MPTP :         | 1-méthyl 4-phényl 1,2,3,6-tétrahydro pyridine                                |
| MS :           | Spectrométrie de masse   |
| MS/MS :        | Spectrométrie de Masse en Tandem   |
| mTOR :         | Mammalian target of rapamycin  |
| Mut :          | Mutation   |
| Myc :          | Myelocytomatosis oncogene  |
| N-terminal :   | Terminaison amine  |
| Nat1 :         | N-acyltransférase 1  |
| NCBI :         | National Center for Biotechnology Information                                |

|                    |   |
|--------------------|---|
| Neo :              | Néomycine   |
| NLS :              | Nuclear localisation sequence   |
| NMDA :             | N-methyl-D-aspartic acid  |
| PA :               | Acide phosphatidique  |
| PKA :              | cAMP-dependent protein kinase   |
| PC :               | Phosphatidylcholine   |
| PCR :              | Polymerase Chain Reaction   |
| PE :               | Phosphatidyléthanolamine  |
| PG :               | Phosphatidylglycérol  |
| PI :               | Phosphatidylinositol  |
| Pfam :             | Protein families database   |
| PGK :              | Promoteur du gène codant la phosphoglycérate kinase1                          |
| PIR :              | Protein Information Resource  |
| PLA <sub>2</sub> : | Phospholipase A <sub>2</sub>  |
| ppm :              | Partie par million  |
| PrESTs :           | Protein Epitope Signature Tags  |
| PS :               | Phosphatidylsérine  |
| Q-trap :           | Quadrupole linear ion trap  |
| qPCR :             | real-time Quantitative Polymerase Chain Reaction                              |
| qRT-PCR :          | real-time Quantitative Reverse Transcription Polymerase Chain Reaction        |
| RMN :              | Résonance magnétique nucléaire  |
| RT-PCR :           | Reverse Transcription Polymerase Chain Reaction                               |
| rtTA :             | Reverse tetracycline-controlled transactivator                                |
| SADE :             | SAGE Adaptation for Downsized Extracts  |
| SAGE :             | Serial analysis of gene expression  |
| SAP102 :           | Synapse-associated protein 102  |
| SELDI-TOF-MS :     | Surface Enhanced Laser Desorption Ionization-Time Of Flight-Mass Spectrometry |
| scFv :             | Single-chain variable fragment  |
| shRNA :            | Short hairpin Ribonucleic Acid  |
| SIB :              | Swiss Institute of Bioinformatics   |
| siRNA :            | Small interfering ribonucleic acid  |
| SMART :            | Simple Modular Architecture Research Tool                                     |

|                      |   |
|----------------------|---|
| sn (sn1, sn2, sn3) : | stereospecific numbered                                     |
| SNP :                | Single nucleotide polymorphism                              |
| STRING :             | Search Tool for the Retrieval of Interacting Genes/Proteins |
| SV40 :               | Simian virus 40 early promoter                              |
| SynDIG1 :            | Synapse differentiation inducing 1                          |
| TAG :                | Triacylglycérol   |
| Taq :                | Polymérase issue de <i>Thermophilus aquaticus</i>           |
| TaqMan :             | Taq Polymerase + PacMan (jeu vidéo) = TaqMan                |
| TE :                 | Enzyme d'étiquetage   |
| TEP :                | Tomographie à Emission de Positons                          |
| Tet :                | Tétracycline  |
| tetO :               | Opérateur Tet   |
| TH :                 | Tyrosine hydroxylase  |
| TRE :                | Tetracycline responsive element                             |
| tTA :                | Tetracycline-controlled transactivator                      |
| TOF :                | Time of flight  |
| UAS :                | Upstream Activating System                                  |
| UHDRS :              | Unified Huntington Disease Rating Scale                     |
| Uniprot :            | Universal Protein resource                                  |
| UTR :                | UnTranslated Region   |
| V <sub>H</sub> :     | Heavy chain variable region                                 |
| V <sub>L</sub> :     | Light chain variable region                                 |
| WT :                 | Wild type   |
| XCMS :               | Various forms (X) of chromatography mass spectrometry       |
| Xe :                 | Xénon   |
| X-Gal :              | Dérivé (X) du galactose                                     |

## I. Contexte au sein du laboratoire

La compréhension du fonctionnement cérébral et l'application clinique de ces connaissances représentent un fort espoir pour de nombreux patients souffrant de maladies neurodégénératives. De nos jours la génomique fonctionnelle possède des outils puissants qui permettent de mieux comprendre le fonctionnement du cerveau.

Avant mon arrivée en thèse, une analyse transcriptomique à grande échelle de différentes régions du cerveau de souris a été effectuée au sein de notre laboratoire. Pour ce faire, la technique SADE (*SAGE Adaptation for Downsized Extracts*) (Virlon et al., 1999) mise au point dans le laboratoire et issue de la technique SAGE (*Serial Analysis of Gene Expression*) (Velculescu et al., 1995) a été employée. Une étude comparative de l'expression des gènes dans plus de dix régions du cerveau de souris a permis d'identifier des gènes principalement exprimés dans chacune de ces régions (marqueurs régionaux) (Brochier et al., 2008). Au total, plus de 300 marqueurs ont été sélectionnés. Certains sont connus de la littérature et d'autres encore non caractérisés. Parmi ceux-ci, environ 60% ont un profil d'expression conservé entre la souris et l'homme. Ces derniers peuvent se montrer intéressants quant à leur participation éventuelle au développement de maladies neurodégénératives ou pour la compréhension des fonctions spécifiques à chaque région du cerveau.

Au cours de ma thèse, je me suis focalisée sur l'étude fonctionnelle de deux de ces gènes dont l'expression est enrichie dans des régions spécifiquement touchées par des pathologies neurodégénératives - *Agpat4*, un marqueur de l'aire tegmentaire ventrale et de la substance noire compacte, dont les neurones dopaminergiques sont sélectivement atteints lors de la maladie de Parkinson ; et *Capucine*, un marqueur du striatum (de Chaldée et al., 2003 et de Chaldée et al., 2006), structure qui dégénère au cours de la maladie de Huntington.

Mais qu'est-ce qu'un biomarqueur, pourquoi est-il important de l'étudier et en quoi sa caractérisation peut-elle être utile à la recherche fondamentale et à la recherche clinique ?

## II. Introduction générale

### 1. Biomarqueurs

Un biomarqueur, ou marqueur biologique, est en général un indicateur spécifique d'un état biologique. Il existe différents types de biomarqueurs : il peut s'agir de paramètres anatomiques ou physiologiques, ou de substances biochimiques ou moléculaires, discutées ici. Selon la définition établie lors d'un atelier conduit sous la responsabilité du *National Institute of Health* en 1998, un biomarqueur désigne "une caractéristique mesurée objectivement (c'est-à-dire avec une précision et une reproductibilité suffisantes) et évaluée comme indicateur :

- de processus physiologiques normaux
- de processus physiologiques pathologiques,
- de l'action des médicaments".

Dans le corps humain, un biomarqueur moléculaire, ou biochimique, peut être détecté dans un fluide biologique (sang, urine ou fluide cérébro-spinal) ou dans un tissu, *in situ* ou après prélèvement.

Actuellement les biomarqueurs sont utilisés dans de nombreux domaines scientifiques, tels que la médecine, la pharmacologie, l'épidémiologie ou la toxicologie. En médecine, un biomarqueur peut servir comme outil pour examiner la fonction des organes ou bien comme moyen diagnostique ou pronostique. Il peut se présenter sous forme de substance introduite dans l'organisme, par exemple le chlorure de rubidium  $^{82}\text{Rb}$ , isotope radioactif utilisé pour évaluer la perfusion du muscle cardiaque par l'imagerie dynamique TEP (Tomographie à Emission de Positron) (deKemp et al., 2000). Il peut aussi être une molécule endogène ; par exemple, la présence d'un anticorps peut indiquer une infection. Plus précisément, un biomarqueur indique un changement qui est en corrélation avec le risque ou la progression d'une maladie, ou la réponse de la maladie à un traitement donné.

Les biomarqueurs sont également utilisés en épidémiologie et en toxicologie pour indiquer l'exposition à diverses substances de l'environnement. Dans ce cas, le biomarqueur peut être la substance externe elle-même (le cadmium, absorbé en fumant du tabac, peut être mesuré directement dans la salive et l'urine (Talio et al. 2010)), ou résulter de la réponse de l'organisme à cette substance (par exemple, des marqueurs protéiques de changements pathologiques précoces ont été décrits dans le sérum de personnes chroniquement exposées à des métaux lourds tels que l'arsenic, le cadmium et le plomb (Kossowksa et al., 2010)). En

pharmacologie, les biomarqueurs permettent d'étudier la toxicité ou l'efficacité d'un médicament par un simple dosage de molécules au lieu d'avoir à observer des symptômes complexes. Deux biomarqueurs corrélés avec les effets toxiques d'un immunoconjugué (MLN8866) utilisé contre le cancer des ovaires ont ainsi été identifiés dans le sérum de singe et la dose maximale tolérée a pu être prédite (Hsieh et al., 2007).

En recherche fondamentale, les marqueurs moléculaires sont utilisés pour identifier différents tissus ou types cellulaires et permettre leur caractérisation et leur étude fonctionnelle. En biologie cellulaire, un biomarqueur est une molécule qui permet l'identification d'un type cellulaire particulier. Cela peut être intéressant surtout dans le cas de cellules morphologiquement trop semblables pour être distinguées, mais qui ont des fonctions différentes. Grâce à ces marqueurs, on peut effectuer un tri cellulaire. Chez la souris, cette approche est basée sur la construction de lignées knock-in où le gène codant le marqueur est remplacé par le gène de la GFP (*Green Fluorescent Protein*). Celle-ci va donc être exprimée sous le contrôle des séquences régulatrices du gène du marqueur. Grâce à la fluorescence émise par la GFP, on peut alors distinguer le type de cellules ciblé et effectuer leur purification. Plus de 1100 lignées knock-in GFP sont disponibles grâce au projet GENSAT ([http://www.gensat.org/MMRRC\\_report.jsp](http://www.gensat.org/MMRRC_report.jsp)).

Deux techniques utilisant des souris knock-in peuvent être employées pour isoler des types cellulaires déterminés : la microdissection par laser et le FACS (*Fluorescence-Activated Cell Sorting*). La microdissection par laser a par exemple été utilisée pour isoler les neurones pyramidaux de la couche V dans le néocortex de souris exprimant l'EYFP (*Enhanced Yellow Fluorescent Protein*) sous le contrôle des séquences régulatrices du gène Thy-1, qui code un marqueur neuronal. Les transcriptomes de ces neurones ont ensuite été comparés entre les cortex moteur et somesthésique (Rossner et al., 2006). La méthode FACS a notamment été utilisée pour purifier les deux principaux sous-types de neurones de projection striataux (Lobo, et al., 2006). Pour ce faire, les auteurs ont eu recours à des souris exprimant l'EGFP (*Enhanced Green Fluorescent Protein*) sous le contrôle des éléments régulateurs du gène d'un marqueur de chaque type cellulaire (le récepteur dopaminergique D1 (*Drd1a*)- pour les neurones striato-nigraux ; le récepteur dopaminergique D2 (*Drd2a*)- pour les neurones striato-palidaux). Les transcriptomes de ces deux populations de neurones ont ensuite été analysés.

Une nouvelle méthode de purification des ARN polysomiaux provenant de types cellulaires déterminés a été développée. Elle consiste à produire des souris transgéniques exprimant la protéine ribosomique L10a en fusion avec l'EGFP, sous le contrôle du

promoteur du gène d'un marqueur spécifique d'un type cellulaire (Heiman et al., 2008). Grâce à des anticorps spécifiques de l'EGFP, les polysomes peuvent être purifiés. Les ARN polysomaux sont ensuite extraits puis analysés par des puces à ADN. Cette technique peut être utilisée pour identifier les transcrits enrichis dans tout type cellulaire génétiquement défini. Grâce à cette méthode, une ressource de centaines d'ARN messagers cellule-spécifiques a été générée pour 24 populations différentes de cellules du système nerveux central (Doyle et al., 2008).

### **1.1. Différentes approches technologiques pour la découverte de nouveaux biomarqueurs moléculaires**

La découverte de nouveaux biomarqueurs est souvent liée au développement d'une nouvelle technologie (Allinson et Brooks, 2004). Les techniques de détection à grande échelle et à haut débit des biomarqueurs moléculaires sont généralement basées sur l'utilisation de biopuces ou du séquençage et sur le criblage différentiel de l'ADN, de l'ARN, des protéines ou des métabolites contenus dans un échantillon biologique (un tissu, un fluide biologique, ...). En particulier, les technologies de séquençage de nouvelle génération (*next generation sequencing*) (Metzker, 2010) et leurs applications (*seq-based methods*), qui permettent la production d'un volume considérable de données à un prix relativement bas, ont révolutionné la recherche scientifique et depuis, la découverte des biomarqueurs moléculaires a pris une nouvelle dimension.

Nous pouvons ainsi distinguer plusieurs types de biomarqueurs :

- génomiques : principalement les SNPs (*Single Nucleotide Polymorphisms*)
- épigénomiques : état de méthylation de l'ADN, accessibilité de la chromatine, modifications des histones
- transcriptomiques : profils d'expression des transcrits
- protéomiques : profils d'expression des protéines ; l'histoire des protéines en tant que biomarqueurs est la plus ancienne
- métabolomiques : catégorie la moins développée, mais actuellement en plein essor.

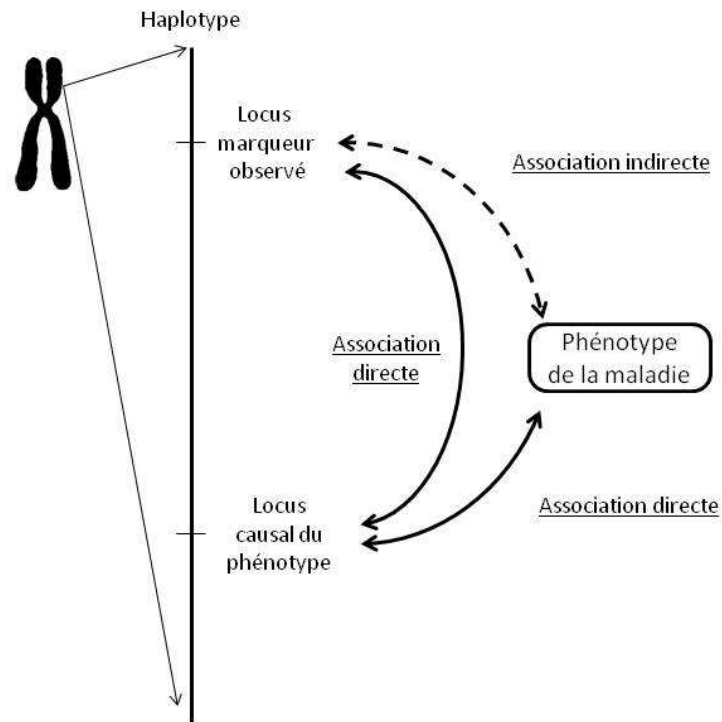


### **1.1.1. L'approche génomique**

En génomique, un biomarqueur est une séquence d'ADN polymorphe, facilement détectable, dont un allèle est associé à un phénotype spécifique ou à la susceptibilité à une maladie. Les biomarqueurs génomiques les plus couramment utilisés sont les SNP. Il s'agit de polymorphismes d'un seul nucléotide qui peuvent se trouver aussi bien dans les régions codantes que non codantes des gènes ou dans les régions intergéniques. Les SNP sont également utilisés en génétique des populations, où ils contribuent à la construction des arbres phylogénétiques.

Les biomarqueurs génomiques peuvent être mis en évidence directement par les techniques de séquençage. En fait, beaucoup d'entre eux sont déjà bien documentés. Le projet international HapMap fournit des informations sur les fréquences alléliques (de plus d'un million de SNPs) dans différentes populations, et sur le déséquilibre de liaison entre les SNP (The International HapMap consortium, 2005). Les données sur les SNP ont permis de concevoir des puces à SNP, une technique basée sur l'hybridation de l'ADN. Le principe des puces à SNP consiste à fixer sur un support en verre des sondes de 25 nucléotides qui portent un SNP en leur milieu. Sur une puce, il y a de nombreuses sondes portant des SNP différents, chacune représentée par plusieurs copies identiques, ce qui permet de rendre le résultat fiable. Une fois la puce préparée, elle peut être hybridée avec l'échantillon d'ADN que l'on souhaite étudier. L'ADN analysé est fragmenté et biotinylé sur l'extrémité avant d'être mis en contact avec la puce. Les fragments complémentaires aux sondes situées sur la puce vont alors s'hybrider. Grâce à la biotine, des molécules de streptavidine fluorescentes peuvent se fixer sur les ADN hybridés sur la puce. La puce est ensuite scannée pour identification des allèles présents dans l'échantillon d'ADN analysé.

Les puces à SNP, le projet international HapMap, qui documente le profil de variation du génome entier, et un nombre important d'échantillons cliniques, pour de nombreuses maladies communes bien caractérisées, ont permis de réaliser des études d'association sur l'ensemble du génome (Figure 1).



**Figure 1 : Principe de base des GWAS (genome wide association studies).**

Pour effectuer une étude GWAS on compare les variations génomiques entre un groupe de « cas » et un groupe de témoins. Les allèles significativement sur-représentés chez les « cas » sont associés au locus causal du phénotype observé.

Les premières études d'association ont été effectuées au milieu des années 2000. Relativement peu de SNP (11560 SNP par puce) étaient placés sur chaque puce et les échantillons étaient regroupés pour diminuer le nombre de puces (Butcher et al., 2005). Depuis 2007, les GWAS sont utilisées pour l'analyse à grande échelle des facteurs génétiques associés à des maladies humaines (The Wellcome Trust Case Control Consortium, 2007). Cette approche consiste en l'étude de la variabilité du génome entier, dans une population d'individus donnée. Dans une population britannique composée d'environ deux mille individus examinés pour sept maladies majeures, et une série de trois mille contrôles, une étude GWAS (The Wellcome Trust Case Control Consortium, 2007) a été effectuée. La comparaison cas-témoins a permis d'identifier 24 associations indépendantes : une pour les troubles bipolaires, une pour la maladie coronarienne, neuf pour la maladie de Crohn, trois pour la polyarthrite rhumatoïde, sept pour le diabète de type 1 et trois pour le diabète de type 2. Chaque échantillon d'ADN a été hybridé sur deux puces, chacune étant capable de génotyper en moyenne 250.000 SNP. Depuis, les études GWAS ont été élargies à plus d'un millier de phénotypes différents (Manolio, 2010).

### 1.1.2. L'approche épigénomique

Malgré le succès des GWAS dans l'identification de *loci* associés à des phénotypes spécifiques (le plus fréquemment des maladies humaines), des études complémentaires s'imposent pour mieux comprendre leur manifestation. L'épigénétique, focalisée sur les modifications de la chromatine, permet d'appréhender les effets de l'environnement et les mécanismes de l'hérédité non basée sur l'ADN.

La chromatine est une structure complexe formée par l'ADN associé à des protéines (principalement les histones). L'ADN comme les histones peuvent subir des modifications, qui peuvent avoir un impact sur l'expression génique. L'ADN eucaryote peut être méthylé sur les cytosines dans les dimères C-G, ce qui est généralement la marque d'une chromatine transcriptionnellement silencieuse. Les histones peuvent être méthylées ou acétylées sur leurs queues riches en lysines. L'état hyperacétylé des histones correspond généralement aux régions transcriptionnellement actives. L'état de la chromatine peut être modifié par le déplacement des histones, par la liaison de facteurs de transcription ou par des enzymes de modification de la chromatine comme méthyltransférases de l'ADN (Dnmt) et les histonedésacétylases (HDAC) (Griffiths et Gore, 2008). La structure de chromatine constitue pour la cellule une mémoire à long terme (Koche et al., 2011). Des changements à grande échelle dans la méthylation de l'ADN et des histones ont notamment été observés dans plusieurs cancers (Jones et Baylin, 2007). Un exemple d'hérédité basée sur la méthylation de l'ADN est la couleur de pelage agouti chez la souris. En effet celle-ci dépend du niveau de méthylation au locus A du gène *Agouti*. A l'état homozygote récessif *aa*, la couleur du pelage est noire, tandis qu'à l'état dominant *AA*, elle est jaune. Les hétérozygotes *Aa* peuvent avoir une variété de couleurs entre jaune et agouti. Plus le locus A est méthylé, plus la couleur est proche de l'agouti. En outre cette marque épigénétique persiste dans la lignée germinale femelle ce qui peut expliquer une hérédité maternelle (Morgan et al., 1999).

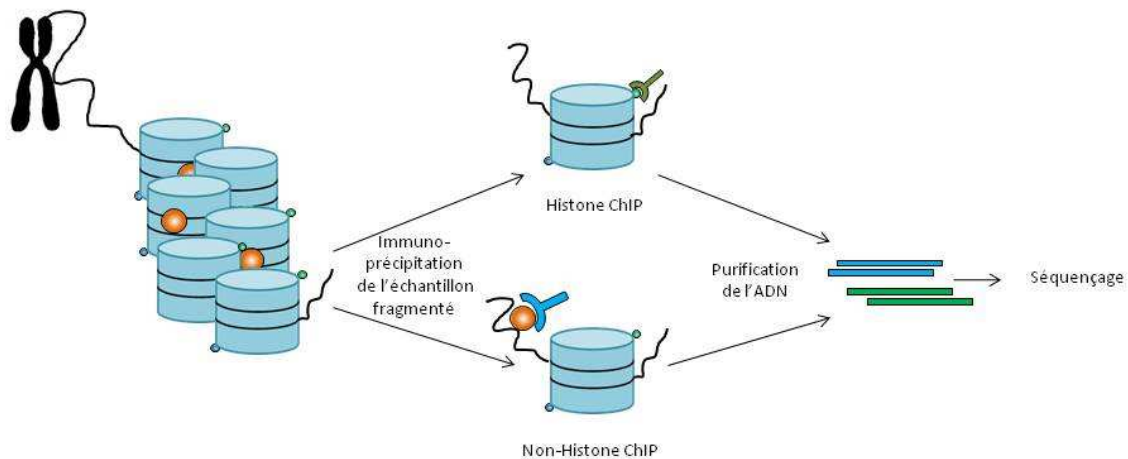
Le développement des techniques de séquençage de nouvelle génération a permis d'étudier les mécanismes épigénétiques et d'analyser les variations interindividuelles de marques épigénétiques à grande échelle. Une des technologies habituellement utilisées pour étudier l'état de méthylation de l'ADN est le bisulfite-seq. C'est une méthode basée sur l'utilisation du bisulfite, qui convertit les résidus de cytosine en uracile, mais laisse des résidus 5-méthylcytosine intacts. L'accessibilité de la chromatine peut être évaluée par la sensibilité à la DNase I. Le plus souvent, les sites hypersensibles à la DNase I se trouvent

dans les régions de la chromatine qui contiennent un gène exprimé. Ils sont générés à la suite d'une liaison de facteurs de transcription qui déplacent les histones.

La technique la plus souvent utilisée pour l'analyse des marques épigénétiques est l'immunoprécipitation de la chromatine (ChIP), qui aujourd'hui peut être suivie par un séquençage (*chromatin immunoprecipitation coupled to massively parallel sequencing* ChIP-seq) (Figure 2). Cette technique a été employée pour la première analyse à grande échelle de marques épigénétiques en 2007 en utilisant des anticorps dirigés contre des histones modifiées. Des cartes de distribution de vingt modifications d'histones sur le génome humain ont ainsi été générées (Barski et al., 2007). La méthylation des résidus cytosine sur le carbone 5 (5<sup>me</sup> C) de l'ADN dans le contexte de séquence CpG est une marque épigénétique commune chez de nombreux eucaryotes, qui peut être détectée en utilisant des anticorps spécifiques contre le 5<sup>me</sup>C, ou en utilisant des protéines qui se lient spécifiquement à l'ADN méthylé (Bock et al., 2010). L'analyse des séquences de régulation, qui contrôlent l'expression spatio-temporelle des gènes, est aussi possible grâce au ChIP-seq. Plusieurs centaines de régions « *enhancers* » de la transcription ont ainsi été identifiées dans le prosencéphale, le mésencéphale et les membres chez les embryons de souris en précipitant la chromatine avec un anticorps spécifique contre p300, une protéine de liaison aux enhancers actifs exprimée de façon ubiquitaire (Visel et al., 2009).

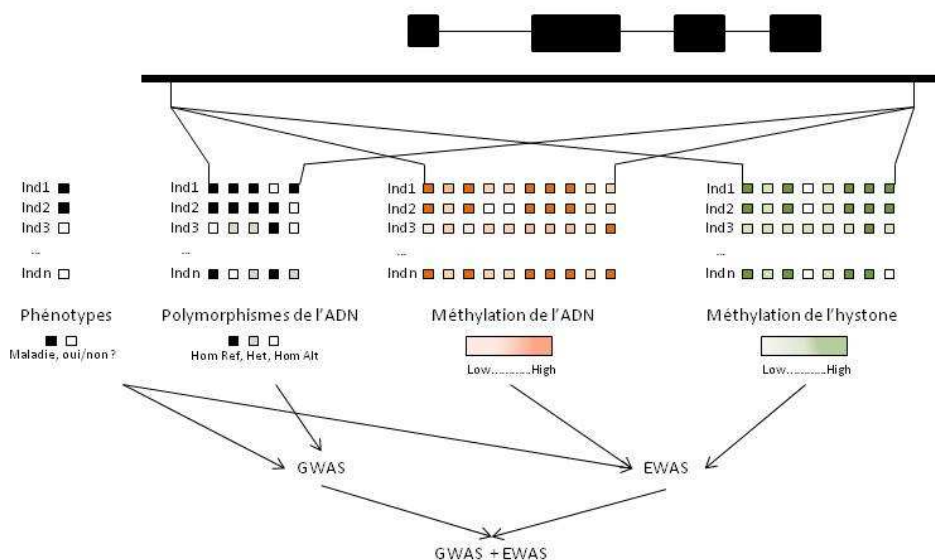
La méthode ChIP-seq a de nombreux avantages par rapport au ChIP-chip, où les fragments obtenus à partir de l'immunoprécipitation de la chromatine sont identifiés par hybridation sur une micropuce. Le ChIP-seq présente en effet une meilleure résolution, moins d'artefacts, une couverture plus grande et une gamme de détection plus large.

Pour comprendre comment des facteurs épigénétiques peuvent être corrélés avec un phénotype, des études d'association à grande échelle de marques épigénétiques (EWAS : *epigenome-wide association studies*) ont été développées. Or le fait que l'état de la chromatine soit susceptible de varier dans le temps, ainsi que d'une cellule à l'autre, pose un problème d'interprétation des résultats, qui pourrait être résolu en combinant GWAS et EWAS (Figure 3). Dans une étude récente combinant les analyses génomiques et épigénomiques, le niveau de méthylation de l'ADN a été déterminé à différents loci de susceptibilité au diabète de type 2. Au locus FTO (*Fat mass and obesity-associated gene*), le niveau de méthylation est significativement plus élevé chez les porteurs de l'haplotype associé au diabète de type 2. L'examen de la séquence a révélé que les allèles des SNPs qui constituent cet haplotype génèrent trois sites CpG supplémentaires (Bell et al., 2010).



**Figure 2 : Représentation schématique de la méthode ChIP-seq.**

Les régions de la chromatine qui interagissent avec des facteurs de transcription ou d'autres protéines associées à la chromatine (*non-histone ChIP*) ou qui contiennent des nucléosomes modifiés (*histone ChIP*) sont immunoprécipités par des anticorps spécifiques. L'ADN immunoprécipité est ensuite purifié et séquencé par une des techniques de séquençage de nouvelle génération.



**Figure 3 : Illustration de la méthode qui combine les résultats des GWAS et EWAS.**

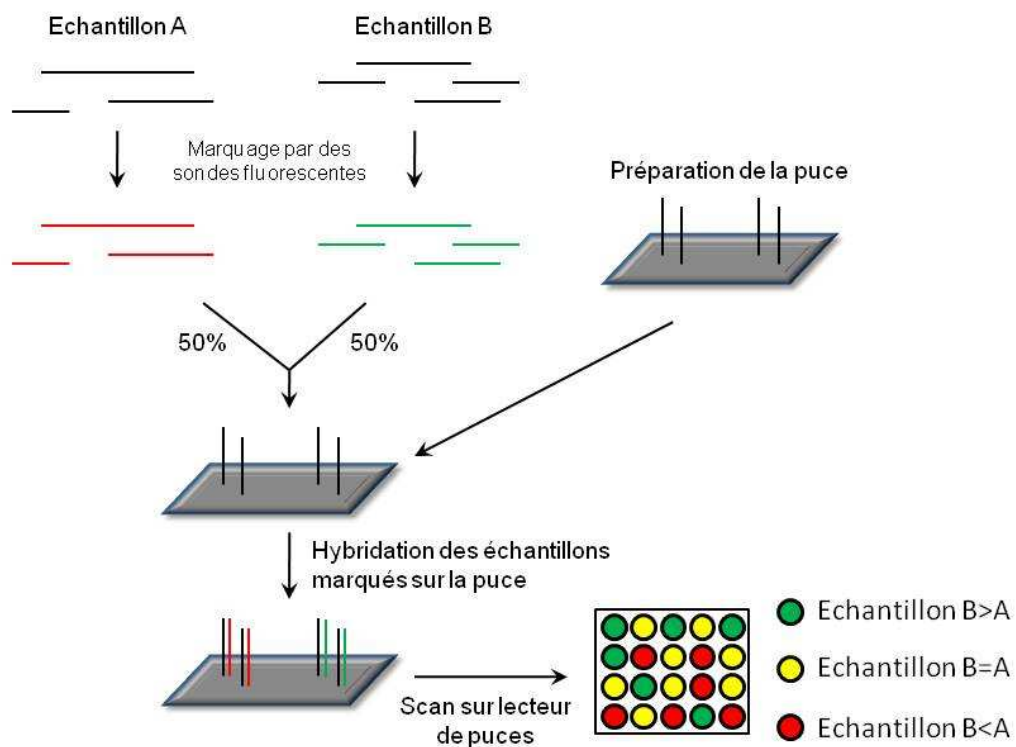
Les GWAS associent la présence d'un allèle à un phénotype. Les EWAS associent un état chromatinien au même phénotype. Deux marques épigénétiques sont prises en compte dans cette étude : le niveau de méthylation de l'ADN et le niveau de méthylation des histones. En combinant les résultats de ces deux études (GWAS+EWAS), effectuées sur les mêmes individus, on peut obtenir une vue plus précise de la corrélation entre les événements moléculaires et leurs effets phénotypiques (Birney, 2011)

### **1.1.3. L'approche transcriptomique**

Chaque gène a son propre champ de profils d'expression, par exemple dans quels tissus et à quel moment il est exprimé, de quelle façon il répond à des perturbations variées. La compréhension du transcriptome est essentielle pour pouvoir interpréter les éléments fonctionnels du génome, révéler la composition moléculaire des cellules et des tissus, mais aussi pour la compréhension du développement et des maladies. Le but de l'approche transcriptomique est de quantifier le changement de niveau d'expression dans différentes conditions ou selon l'état physiologique de l'échantillon examiné, et d'analyser les profils d'épissage.

Les techniques utilisées dans l'approche transcriptomique sont fondées d'une part sur l'hybridation et de l'autre part sur le séquençage. Dans le premier cas, des échantillons d'ARN marqués sont hybridés à des séquences complémentaires des transcrits d'intérêt fixées sur une puce (Figure 4). Cette technique permet d'établir des catalogues de gènes spécifiquement exprimés dans certaines cellules ou tissus, donc des marqueurs cellulaires ou tissulaires (Liang et al, 2006). On peut aussi analyser l'expression des gènes dans différentes conditions ou à différents stades de développement et caractériser ainsi la dynamique d'expression de chaque gène du génome (Shi et al., 2008). Dans le cas de pathologies, on peut comparer les différences d'expression génique entre les tissus sain et malade pour trouver des gènes éventuellement impliqués dans le processus de la maladie (Tseveleki et al., 2010). Une limitation de cette approche est le niveau du bruit de fond qui peut être élevé par suite d'hybridations croisées. La gamme dynamique, limitée par le bruit de fond et la saturation des signaux, est également restreinte. En outre, la comparaison des niveaux d'expression entre différentes expériences est difficilement réalisable et nécessite des méthodes de normalisation complexes.

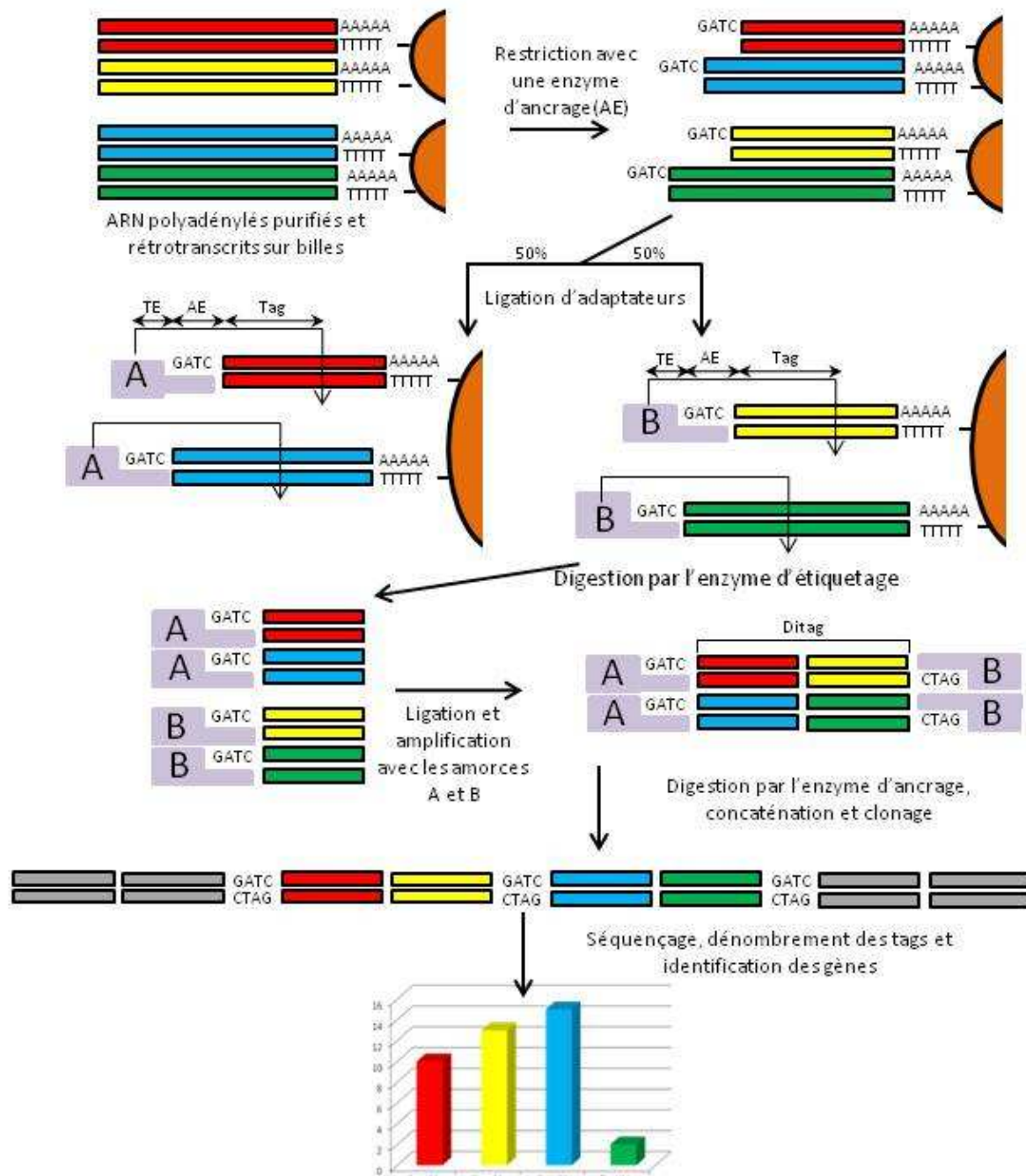
Contrairement aux techniques basées sur l'hybridation, les techniques basées sur le séquençage ne nécessitent pas la connaissance préalable de la séquence des transcrits. Parmi les techniques développées avant l'essor du séquençage à haut débit, je ne vais mentionner que la technique SAGE, qui a permis l'identification des deux marqueurs génétiques que j'ai étudiés au cours de ma thèse. Elle est basée sur la production de courtes étiquettes gènes-spécifiques et leur séquençage (Figure 5) (Velculescu et al., 1995, Saha et al., 2002).



**Figure 4 : Le principe de la technique des puces à ADN.**

Les puces à ADN sont des supports solides sur lesquels sont disposées de manière ordonnée des molécules d'ADN complémentaires des transcrits que l'on souhaite analyser. La fabrication de la puce à ADN peut se faire soit par dépôt d'une banque d'ADNc amplifiés par PCR, soit par synthèse *in situ* d'oligonucléotides de 20 à 25 nucléotides sur le support. Une vingtaine d'oligonucléotides est utilisée par transcrit pour assurer la fiabilité d'hybridation. Les ARN sont extraits des échantillons biologiques puis marqués par des fluorophores spécifiques de chaque échantillon et hybridés sur les puces. L'intensité de chaque fluorophore est mesurée afin d'identifier les gènes exprimés dans les différents échantillons.





**Figure 5 : Schéma de la technique SAGE utilisée au laboratoire.**

Les ARN polyadénylés sont purifiés à partir de l'échantillon biologique grâce à des oligonucléotides poly (dT) fixés sur billes magnétiques puis ils sont rétrotranscrits en ADNc. Après digestion par une enzyme d'ancrage (AE) à site de coupure fréquent (ici : Sau3aI), l'échantillon est séparé en deux lots égaux. Chaque lot est ensuite ligaturé avec un des deux adaptateurs A et B contenant un site de reconnaissance pour une enzyme de type II (enzyme d'étiquetage - TE), qui coupe à une distance définie de son site de reconnaissance donnant ainsi naissance à un tag. Les ditags sont ensuite générés par ligation puis amplifiés par PCR et concaténés. Le séquençage des concatémères permet d'associer les tags aux transcrits dont ils proviennent. En dénombrant les tags on peut établir le profil d'expression des gènes dans l'échantillon étudié.



Malgré leur petite taille, chacune de ces étiquettes contient en théorie une information suffisante pour identifier sans ambiguïté le transcrit dont elle provient. Le séquençage en série de plusieurs étiquettes concaténées au sein du même clone augmente l'efficacité de l'analyse. A l'aide du séquençage on peut identifier les différentes étiquettes et leur dénombrement permet d'établir le niveau d'expression des transcrits dont elles proviennent dans les tissus analysés. Grâce à une miniaturisation de cette technique (Virlon et al., 1999), plusieurs marqueurs de différentes régions du cerveau de souris, dont certaines sont de petite taille (par exemple l'aire tegmentaire ventrale : deux prélèvements de 0,07mm<sup>3</sup> par souris), ont pu être mis en évidence au laboratoire (de Chaldée et al., 2003 ; Brochier et al. 2008).

L'avènement du séquençage à haut débit de nouvelle génération (*Next Generation Sequencing*) a révolutionné l'étude du transcriptome en permettant l'analyse de l'ARN par le séquençage massif de l'ADNc. Cette méthode appelé RNA-seq a de nombreuses avantages par rapport aux techniques précédentes : elle n'est pas limitée par la connaissance préalable de la séquence des transcrits, elle permet de localiser facilement les frontières de la transcription et les jonctions intron-exon à un nucléotide près, et, grâce à sa résolution, elle est capable d'identifier des variations au sein d'une séquence tels que les SNPs dans les régions codantes (Cloonan et al., 2008). Un autre avantage du RNA-seq est qu'il ne génère quasiment pas de bruit de fond et qu'il permet une quantification précise.

#### **1.1.4. Approche protéomique**

En plus des études génomiques et transcriptomiques, différents biomarqueurs peuvent être identifiés par des méthodes de protéomique. Contrairement au génome, le protéome est beaucoup plus dynamique et les profils d'expression protéique des cellules dans un organisme varient en fonction de son état physiologique, du stade de développement ou de la différenciation cellulaire et des facteurs environnementaux. La complexité du protéome augmente aussi à cause de l'épissage alternatif et à cause des modifications post-traductionnelles (notamment la phosphorylation, la glycosylation ou l'ubiquitinylation). De nombreuses techniques ont été développées dans le but d'identifier des biomarqueurs protéomiques. Les puces à protéines (*protein microarray*) où une banque d'anticorps ou de protéines et immobilisée à haute densité spatiale sur une lame de microscope sont largement utilisées. Les anticorps fixés par une liaison covalente sont sondés avec les extraits protéiques de l'échantillon analysé. Un deuxième anticorps biotinylé est ensuite appliqué et se lie sur les protéines capturées, formant un complexe immun. Les seconds anticorps sont ensuite détectés

avec un anticorps antibiotine universel conjugué à un fluorophore (Tabibiazar et al., 2006). Les techniques de gel d'électrophorèse à 2 dimensions (bidimensionnelle) permettent de construire, avec une grande résolution, la carte des protéines exprimées dans un tissu à un moment donné (Rite et al., 2007). Les techniques MALDI-TOF-MS et SELDI-TOF-MS (Matrix-Assisted et Surface Enhanced Laser Desorption Ionization-Time Of Flight-Mass Spectrometry) associent une étape où les protéines déposées sur une barrette se lient sur des surfaces biochimiquement actives (sites d'ancrage) ou sur la surface d'une puce et une étape de détection par désorption et ionisation par laser suivie d'une analyse par spectrométrie de masse (Araki et al., 2010).

### **1.1.5. Approche métabolomique**

La métabolomique consiste à déterminer de façon quantitative des changements de composition chimique, liés au temps ou à différents stimuli, dans un système biologique intégré (fluide biologique, tissu ou cellules). Les biomarqueurs métabolomiques sont détectés principalement par les techniques de RMN (résonance magnétique nucléaire) et de spectrométrie de masse (MS). L'utilisation de la spectrométrie de masse représente un progrès majeur de ces dernières années en permettant d'obtenir des profils métabolomiques directement à partir d'échantillons de petite taille. Cette approche, employée pour l'étude fonctionnelle d'Agpat4 durant ma thèse, fera l'objet d'une description plus approfondie.

#### **1.1.5.1. Profilage**

La spectrométrie de masse est un outil puissant pour déterminer la structure des molécules organiques (Tabet et Rebuffat, 2003). En effet elle donne accès à la masse moléculaire, ce qui permet d'identifier des molécules recherchées ainsi que des molécules inconnues, après détermination de leur composition élémentaire. Pratiquement un spectromètre de masse (dont le fonctionnement est illustré sur la Figure 6) génère un spectre de masse qui est représenté sous forme d'un diagramme portant le rapport masse/charge en abscisse et l'intensité du signal en ordonnée. L'intensité indiquée est souvent relative, la hauteur des pics étant exprimée en pourcentage de celle du pic le plus intense (pic de base) trouvé dans le spectre de masse. Les limites de détection peuvent être améliorées par le couplage avec des techniques chromatographiques.



**Figure 6 : Schéma général des étapes de fonctionnement d'un spectromètre de masse**

Un spectromètre de masse est composé de trois parties fonctionnelles : une source d'ionisation (sous vide ou à pression atmosphérique) permettant le passage des molécules de l'échantillon sous la forme d'ions à l'état gazeux, un analyseur pour séparer les différents ions produits selon leur rapport masse sur charge ( $m/z$ ), et un détecteur pour mesurer le signal appartenant à chacun des ions. Le signal obtenu est ensuite traité par des systèmes informatiques fournissant un spectre de masse.

#### 1.1.5.1.1. Ionisation

Différentes méthodes d'ionisation sont utilisées en spectrométrie de masse. On peut les diviser en deux groupes : les méthodes directes sont utilisées pour ioniser des composés apolaires, volatiles et résistants à la chaleur utilisée pour amener l'échantillon à l'état gazeux sous vide ; les méthodes indirectes, considérées comme des techniques d'ionisation douces, sont utilisées pour ioniser des molécules fragiles (protéines, oligosaccharides, acides nucléiques). Dans ce cas, le passage à l'état gazeux se fait rapidement et conjointement à l'ionisation.

Parmi les méthodes d'ionisation directes, on peut mentionner :

- l'ionisation électronique (EI) où les molécules de l'échantillon amené à l'état gazeux, sont frappées par un faisceau d'électrons.

- l'ionisation chimique (CI) où la substance à analyser est injectée dans la source avec un large excès de gaz réactif (méthane, ammoniac, isobutane...) ionisé par un faisceau d'électrons. Ce dernier ionise la substance par transfert de charge.

Les méthodes d'ionisation indirectes sont les suivantes:

- le bombardement par atomes rapides (FAB) est une technique où la substance à analyser est dissoute dans une matrice liquide puis bombardée avec une énergie élevée par un faisceau d'atomes accélérés (Ar, Kr ou Xe). La substance à analyser doit de ce fait être soluble dans le liquide servant de matrice, ce qui pose un problème lorsque l'on analyse des mélanges de composés hydrophiles et lipophiles.

- la désorption-ionisation laser assistée par matrice (MALDI) permet d'ioniser un échantillon solide préalablement dispersé sur une matrice en l'irradiant par des photons émis par un laser. Les molécules de la matrice excitées par les photons s'ionisent et transfèrent leur charge aux molécules de l'échantillon.

- La désorption sous ionisation chimique est une ionisation chimique classique, utilisant un chauffage rapide de l'échantillon pour l'amener à l'état gazeux avant qu'il n'ait le temps d'être dégradé

- l'électronébulisation ou électrospray (ESI) est une technique qui permet de désolvater et d'ioniser, à pression atmosphérique, les molécules d'un échantillon dissous dans un solvant sous l'influence d'un champ électrique (Fenn et al. 1989). L'échantillon en solution est introduit dans une chambre à pression atmosphérique par un capillaire fin. A la sortie du capillaire la solution est expulsée en formant de petites gouttelettes. Celles-ci sont évaporées par collision avec un gaz inerte (azote chaud). Un champ électrique provoque la désorption et l'ionisation des composés en phase gazeuse. L'électronébulisation peut facilement être couplée à la chromatographie en phase liquide. Considérée à l'origine comme une source d'ionisation dédiée à l'analyse de polymères, son utilisation s'est rapidement élargie à l'analyse des petites molécules.

#### **1.1.5.1.2. Chromatographie**

L'ionisation des macromolécules est un processus au rendement faible et variable. Or les échantillons étudiés sont souvent complexes, et la plupart des molécules qu'ils contiennent sont susceptibles de s'ioniser, ce qui rend non seulement le spectre de masse plus difficile à analyser, mais diminue en outre le rendement d'ionisation des composés d'intérêt. C'est

pourquoi il est nécessaire de séparer les analytes des autres composants de l'échantillon avant d'analyser leur masse (Ciccimaro et Blair, 2010). Pour cela, une étape de chromatographie précède l'ionisation. Les ionisations EI et CI, qui ont lieu sous vide, sont préférentiellement utilisées en couplage avec la chromatographie en phase gazeuse. En revanche, les méthodes d'ionisation douce à pression atmosphérique, sont principalement utilisées en couplage avec la chromatographie en phase liquide. La chromatographie en phase gazeuse (GC) s'applique principalement aux composés gazeux ou susceptibles d'être vaporisés par chauffage sans décomposition. On peut aussi recourir à une dérivation chimique pour rendre les analytes volatils. Certains types d'échantillons sont particulièrement bien adaptés à une analyse par GC-MS. C'est notamment le cas des terpènes et des huiles essentielles, qui sont volatils et ne s'ionisent pas par la technique LC-MS (*liquid chromatography-mass spectrometry*). La chromatographie en phase liquide ne requiert pas que les métabolites soient volatils. En conséquence, la LC-MS peut analyser un éventail beaucoup plus large d'espèces chimiques que la GC-MS. La LC-MS est donc plus adaptée pour rechercher des métabolites inconnus. La séparation par chromatographie liquide à haut rendement couplée avec l'ESI-MS permet d'obtenir une couverture aussi complète que possible du métabolome. Elle est de ce fait devenue l'une des techniques les plus employées pour la bioanalyse des composés de faible masse moléculaire, et elle est en fort développement pour l'analyse des molécules de taille plus importantes telles que les protéines ou les lipides.

#### **1.1.5.1.3. Analyseurs de masse**

Une fois séparées par chromatographie en phase gazeuse ou liquide (dans le cas de mélanges complexes comme ceux rencontrés en bioanalyse), les molécules ionisées sont triées selon leur rapport masse/charge. Ce tri se fait de manière différente selon l'analyseur utilisé. Il existe deux catégories d'analyseurs : les analyseurs à faisceaux d'ions (par exemple l'analyseur à temps de vol et le quadripôle) et les analyseurs à piégeage d'ions. Dans le système de l'analyseur TOF (*time of flight*) les masses sont séparées sur la base de leur temps de vol. Le temps de vol est le temps que chaque ion met pour arriver au détecteur. Le déplacement des ions vers le détecteur est induit par une différence de potentiel à la sortie de la source d'ions. Plus les composés sont légers, plus leurs temps de vol est court. Le quadripôle est constitué de quatre barres parallèles disposées symétriquement autour d'un axe, associées électriquement deux par deux et créant ainsi un champ électrique d'oscillation. Celui-ci laisse passer certaines masses tandis que d'autres sont sur une trajectoire qui ne leur

permet pas d'atteindre le détecteur. Le quadripôle fonctionne ainsi comme un filtre de masse. Les analyseurs à piégeage d'ions détectent les ions expulsés de façon sélective à une valeur de rapport  $m/z$  donnée d'un champ tridimensionnel de piégeage. Il existe dans cette catégorie des analyseurs à piège linéaire (Q-trap ou LTQ) qui imposent aux ions des allers et retours linéaires au milieu de quatre barres alimentées par des potentiels. Les analyseurs à ultra haute résolution comme l'Orbitrap piègent des paquets d'ions et mesurent des fréquences de leur oscillation.

Chaque analyseur possède des caractéristiques et des performances analytiques qui peuvent être définies par les paramètres suivants : les limites minimale et maximale de  $m/z$  au-delà desquelles l'analyseur ne pourra plus mesurer de signal, la sensibilité qui indique la quantité d'échantillon nécessaire pour détecter un signal (interprétable) dans le spectre de masse, la résolution, qui représente la capacité de l'analyseur à distinguer deux composés de rapports  $m/z$  voisins, et la précision sur la mesure de  $m/z$  qui correspond à la faculté de mesurer une masse la plus proche possible de la masse théorique de la molécule analysée. Cette précision s'exprime le plus souvent en « partie par million » (ppm). Pour l'Orbitrap (utilisé pour mon travail), la limite maximale du rapport  $m/z$  est 4000 et la précision des mesures de  $m/z$  est 1-2 ppm (Makarov et al., 2006).

#### **1.1.5.1.4. Analyse des données**

Les chromatogrammes générés par le spectromètre doivent être convertis en données sous un format compatible avec une analyse informatique. Pour ce faire différents logiciels peuvent être utilisés comme XCMS (Smith et al., 2006) ou Camera (<http://www.bioconductor.org/packages/release/bioc/html/CAMERA.html>). XCMS est un logiciel qui extrait les signaux  $m/z$  des chromatogrammes, regroupe ces signaux et corrige les temps de rétention entre différents chromatogrammes et restitue les données sous forme de matrice contenant des variables, identifiées par un rapport  $m/z$  et un temps de rétention. Camera regroupe les signaux issus d'un même métabolite en calculant la corrélation de la forme des pics à un même temps de rétention et annote les pics isotopiques, adduits et fragments à partir de la matrice générée par XCMS. Cette annotation facilite la précision de la composition élémentaire.

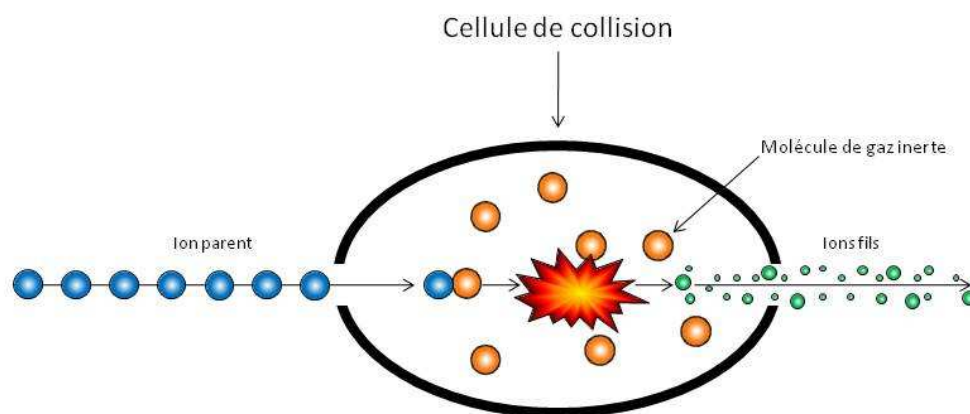
Les données obtenues doivent ensuite être triées pour éliminer les biais dus au traitement des échantillons et à l'analyse. Cette étape de tri repose sur l'analyse de « contrôle-qualité » (blancs expérimentaux, échantillons biologiques dilués en série).

Une partie des signaux obtenus peut être annotée en comparant les  $m/z$  expérimentaux aux masses des molécules enregistrées dans les bases de données. Parmi ces bases de données citons KEGG (*Kyoto Encyclopedia of Genes and Genomes*), HMDB (*The Human Metabolome Database*) et METLIN (*Metabolite and Tandem MS Database*). KEGG est une ressource composée de 16 bases de données, classées en information sur les systèmes, information génomique, et information chimique. Elle contient 17601 métabolites. HMDB est une base de données contenant des informations détaillées sur les métabolites présents dans le corps humain. Elle contient plus de 7900 métabolites hydrosolubles et liposolubles. En outre, environ 7200 protéines (et ADN) en font également partie. METLIN est une base de données de métabolites contenant plus de 42.000 structures.

### **1.1.5.2 Spectrométrie de masse en tandem**

Pour déterminer la structure d'une molécule les informations obtenues en mode de spectrométrie de masse simple peuvent s'avérer insuffisantes. On devra alors recourir à d'autres modes d'acquisition comme le mode de spectrométrie de masse en tandem (MS/MS) ou à fragmentations successives (MS<sup>n</sup>). Dans le mode MS/MS (Figure 7), différentes liaisons dans la molécule cible se rompent selon l'énergie déployée. Les liaisons les plus faibles ont tendance à se dissocier en premier, ce qui produit des ions correspondant à des pertes neutres d'eau, d'ammoniac, ou de dioxyde de carbone. Le mode MS/MS est un moyen de caractériser une molécule complexe, qui n'a pas été reconnue par une simple MS. Des expériences de MS<sup>n</sup> (MS<sup>3</sup>, MS<sup>4</sup>, etc.) peuvent être nécessaires pour élucider la structure complète de la molécule étudiée.

La validation formelle de l'identité d'un métabolite sera finalement effectuée par comparaison des rapports  $m/z$  et des temps de rétention entre le métabolite identifié expérimentalement et le métabolite de référence synthétique.



**Figure 7 : Schéma de la MS/MS.**

La première étape de MS/MS consiste à isoler les ions d'un même rapport  $m/z$  (ion parent) au moyen d'un analyseur, à les fragmenter, puis à analyser les ions produits par fragmentation (ions fils). La majorité des instruments LC-MS avec des capacités de MS/MS emploient une dissociation induite par collision (CID) avec un gaz rare (Ar, He).

### 1.1.5.3 Applications

Un des premiers travaux utilisant la LC-MS et le profilage métabolique a permis d'identifier une nouvelle catégorie de lipides du système nerveux central, les N-acyl taurines, comme substrats endogènes de l'enzyme FAAH (une amide hydrolase des acides gras). Ces lipides sont considérablement plus abondants dans le cerveau et la moelle épinière des souris knock-out pour la FAAH que chez les souris sauvages (Saghatelian et al., 2004). Le profilage métabolique est aussi très utile pour identifier de nouveaux biomarqueurs de maladies. Ainsi, un ensemble de 14 biomarqueurs montrant une excellente discrimination entre les individus souffrant d'onchocercose (une maladie tropicale due à un nématode parasite *Onchocerca volvulus*) et les individus sains a été révélé par une analyse en LC-MS sur des échantillons de sérum et de plasma (Denery et al., 2010). Ces biomarqueurs pourront être utilisés dans un but diagnostique. Une autre étude a été effectuée sur des échantillons de fluide cervico-vaginal humain dans des cas de naissances à terme et de naissances prématurées spontanées pour trouver des biomarqueurs permettant de prédire un risque de naissance prématurée spontanée (Shah et al., 2009). Les protéines présentes dans le fluide cervico-vaginal ont été caractérisées par spectrométrie de masse en tandem couplée à la chromatographie liquide (LC-MS/MS). Trois protéines significativement plus abondantes en cas de naissance prématurée spontanée (l'isoforme 1 de la desmoplakine, la stratifine, et le précurseur 1 de la thrombospondine) ont été proposées en tant que biomarqueurs pronostiques.



## **2. Caractérisation fonctionnelle des biomarqueurs**

Pour comprendre le fonctionnement des cellules, des organismes ou la physiopathologie des maladies, l'identification de marqueurs ne suffit pas et des études fonctionnelles sont nécessaires. Plusieurs approches ont été développées pour caractériser un marqueur de fonction inconnue.

### **2.1. Invalidation/surexpression**

Parmi les stratégies utilisées fréquemment en génétique fonctionnelle, l'invalidation et la surexpression des gènes se sont avérées particulièrement avantageuses. La technique permettant d'inactiver un gène (la technique Knock-Out) est devenue une méthode importante dans la compréhension du rôle d'un gène particulier et des conséquences de son inactivation sur le développement et la physiologie de l'organisme. Son application au modèle murin représente un choix technologique judicieux pour étudier la fonction des gènes chez les mammifères, et pour nous permettre d'extrapoler ces résultats à l'espèce humaine.

#### **2.1.1. Invalidation génique chez la souris**

##### **2.1.1.1. La souris comme organisme modèle**

##### **2.1.1.1.1. Histoire de l'utilisation de la souris en génétique**

La souris a été utilisée pour étudier la génétique des mammifères depuis le début du XX<sup>e</sup> siècle. La première étude génétique de la souris a été effectuée par le biologiste français Lucien Cuénot en 1902 qui, grâce à son étude de l'hérédité de la pigmentation du pelage, a pu vérifier chez la souris les lois de Mendel qui jusqu'alors n'avaient été démontrées que sur les plantes. Quelques années plus tard, aux États-Unis, William Castle et Clarence Little ont commencé une série d'études génétiques sur la souris et, en 1909, C. Little a développé la première souche consanguine de souris de laboratoire, la souris DBA (dilute brown non-agouti). Une souche consanguine, obtenue par des accouplements entre frères et sœurs pendant au moins vingt générations, a un taux d'hétérozygotie proche de zéro. Son uniformité génétique (isogénicité) et phénotypique fait d'elle un modèle stable pour l'étude des conséquences phénotypiques d'une manipulation génétique. Aujourd'hui, il existe plusieurs centaines de souches de souris consanguines, et beaucoup des souches consanguines

originelles sont encore en usage. La souche C57BL/6J, créée par C. Little en 1921, est devenue le premier mammifère après l'homme dont le génome ait été entièrement séquencé (Mouse Genome Sequencing Consortium, 2002). Le progrès rapide des technologies de génétique au cours des années 1980 a conduit à l'élaboration de souris transgéniques et knock-out. Un centre de génétique de la souris comme le *Jackson Laboratory*, fondé par Clarence Little en 1929 aux Etats-Unis, possède aujourd'hui plus de 5000 lignées de souris génétiquement définies.

#### **2.1.1.1.2. Intérêt de ce modèle pour la recherche**

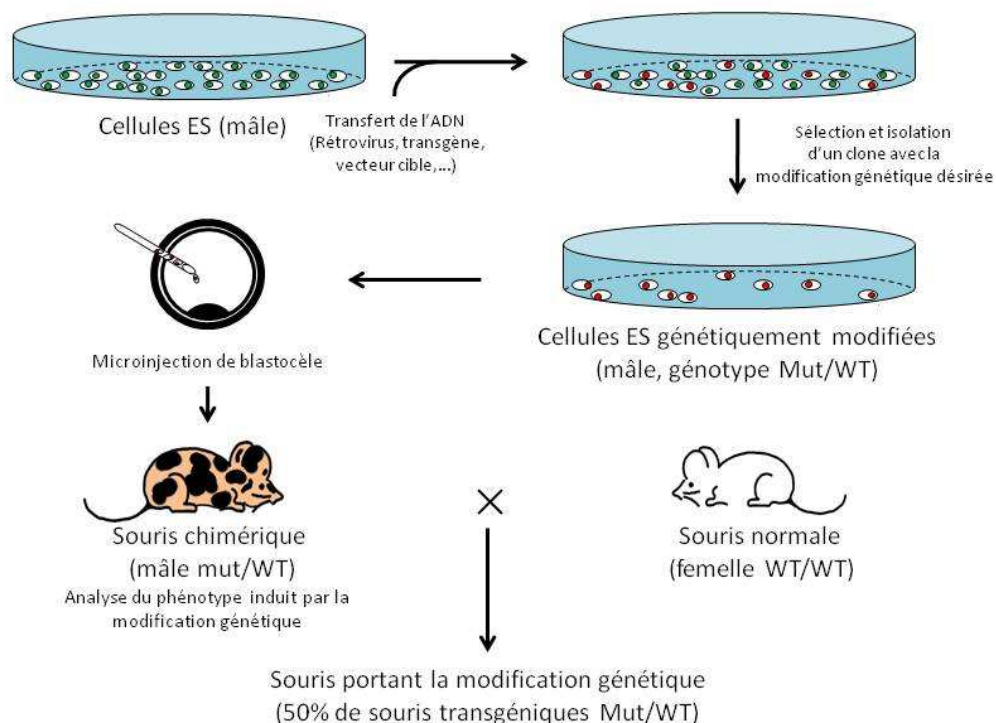
Au siècle dernier, la souris a été choisie comme organisme modèle pour la recherche en génétique des mammifères et les perspectives scientifiques de son utilisation étaient très prometteuses. Bien que la levure, *Caenorhabditis elegans* ou la drosophile puissent être de bons modèles pour certaines études (cycle cellulaire, développement, etc...), la souris est un meilleur modèle pour l'étude des systèmes complexes communs aux mammifères (systèmes immunitaire, neurologique, musculaire, etc...). Par comparaison avec d'autres mammifères, la souris a plusieurs avantages tels qu'un cycle de vie court, une reproduction rapide et une fertilité importante. En outre, grâce à sa petite taille, son élevage est facile et à faible coût. Le fait que cela soit un mammifère avec des caractéristiques génétiques et physiologiques proches de celles de l'homme et le nombre important d'études déjà réalisées sur elle, fait de la souris un modèle de choix pour l'étude de la physiologie et des maladies humaines. L'étude des souris transgéniques permet de comprendre plus précisément l'origine et l'évolution de maladies génétiques et de développer de nouvelles thérapies. Par exemple, les souris déficientes pour la dystrophine et l'utrophine montrent de nombreux signes typiques de la dystrophie musculaire de Duchenne. Leur étude a permis de mettre en évidence que l'utrophine et la dystrophine ont des rôles complémentaires dans le muscle squelettique (Deconinck et al., 1997). Ce modèle est actuellement utilisé pour les essais de nouvelles thérapies (Goyenvallée et al., 2010).

#### **2.1.1.2. Production de souris génétiquement modifiées**

Les premières souris transgéniques ont été produites en infectant des embryons par des rétrovirus ou encore, quelques années plus tard, par micro-injection d'ADN dans le pronucléus d'un zygote. Mais les techniques de génie génétique chez la souris se sont

développées grâce à l'isolement et la mise en culture des cellules souches embryonnaires (ES) (Martin, 1981).

Les cellules souches sont des cellules pluripotentes qui peuvent se diviser de façon illimitée tout en gardant la faculté de se différencier en n'importe quel type de cellules et cela même en culture *in vitro*. Elles peuvent être génétiquement modifiées *in vitro* et ont la capacité de coloniser tous les tissus de l'embryon hôte dans lequel elles sont injectées. Si elles colonisent la lignée germinale, la modification génétique pourra être transmise à la descendance (Figure 8).



**Figure 8 : Représentation schématique de la production de souris génétiquement modifiées.**

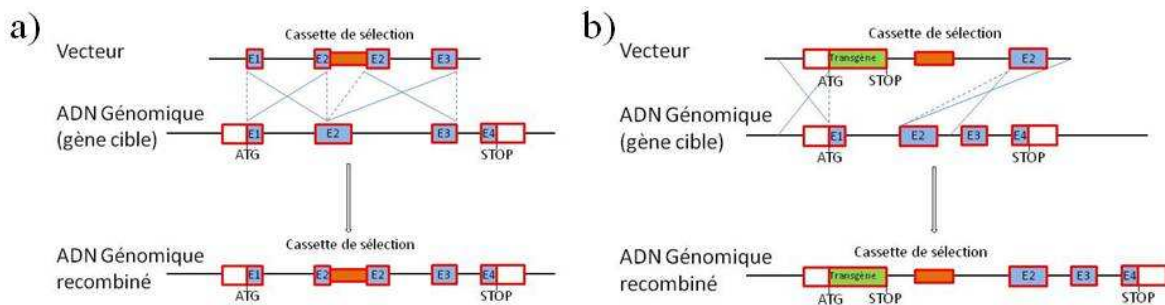
La construction qui porte la modification génétique peut être obtenue par sous-clonage traditionnel, dans un plasmide ou un virus, ou par la technique de *recombineering* qui utilise la recombinaison homologue dans *Escherichia coli* (Liu et al., 2003). Les cellules ES modifiées sont sélectionnées grâce au gène de sélection associé à la modification génétique. Ces cellules seront ensuite multipliées en culture et injectées dans un blastocèle hôte, lequel est réimplanté dans l'utérus d'une femelle porteuse. Le plus souvent les cellules ES et l'embryon hôte sont issus de souches dont le pelage est de couleur différente, ce qui permet d'identifier aisément les souris chimériques, dont une partie des cellules provient de l'embryon hôte, et l'autre provient des cellules ES. Le croisement de ces souris chimériques avec des souris sauvages aboutit, si les cellules ES modifiées ont colonisé la lignée germinale des chimères, à des souris hétérozygotes dont toutes les cellules contiennent un allèle modifié.

### **2.1.1.3. Méthodes d'inactivation génique chez la souris**

#### **2.1.1.3.1. Knock-out et knock-in**

La méthode qui permet la recombinaison homologue entre une molécule d'ADN exogène et une séquence endogène homologue (Figure 9) a été développée chez les eucaryotes supérieurs à la fin des années quatre-vingts (Capecchi, 1989). Cette méthode permet d'inactiver un gène en produisant des mutations ponctuelles, des réarrangements chromosomiques, ou en le délétant. La souris dont un gène a été inactivé est une souris knock-out pour ce gène précis. L'avantage des mutations ponctuelles par rapport à la délétion du gène entier, est qu'elles offrent la possibilité d'analyser la fonction du gène de façon plus subtile, mais cela suppose qu'on ait une idée préalable de la fonction de ce gène ou au moins une hypothèse sur l'effet de ces mutations. Par exemple les souris portant une mutation ponctuelle dans le gène codant la troponine cardiaque I (cTnI) constituent un modèle très utile pour l'étude et la compréhension du développement de la cardiomyopathie restrictive, une maladie qui est due à la même mutation chez l'homme (Du et al., 2006).

Une variante intéressante de la technique de knock-out est la technique de knock-in. Elle consiste à remplacer un gène par une séquence d'ADNc issue d'un autre gène. L'avantage de cette technique est que le transgène reste sous le contrôle du promoteur du gène d'origine. Par exemple un modèle murin de la maladie de Huntington a été généré en remplaçant un fragment du gène de la huntingtine murine par la séquence humaine orthologue mutante (Menalled et al., 2003). Le transgène peut aussi servir de gène rapporteur si son produit est facilement détectable et permettre ainsi d'analyser l'expression du gène qu'il remplace. C'est le cas par exemple du gène LacZ qui code la  $\beta$ -galactosidase, dont l'activité peut être facilement révélée par une simple coloration X-Gal.



**Figure 9 : Principe général de l'inactivation d'un gène par recombinaison homologue.**

a) Dans le cas du knock-out, le vecteur contient une cassette de sélection insérée dans un exon (les exons sont représentés par des rectangles et la séquence codante est en noir) entouré des brins d'homologie qui permettront la recombinaison homologue. Après recombinaison, le gène ciblé est inactivé par l'insertion de la cassette de sélection conduisant à une mutation perte de fonction. b) Dans le cas du knock-in, en plus de la cassette de sélection, un transgène est introduit dans le vecteur. Après la recombinaison homologue, le gène ciblé est inactivé et le transgène placé sous le contrôle du promoteur et des séquences de régulation du gène ciblé est exprimé à sa place.

Il est aussi possible d'inactiver un gène en utilisant des nucléases artificielles à doigts de zinc. Il s'agit d'une technique basée sur une coupure d'ADN double brin et sa réparation. Les nucléases à doigts de zinc comportent un domaine de liaison à l'ADN (chaque nucléase reconnaît une séquence spécifique) associé au domaine nucléaire de l'enzyme de restriction FokI qui n'est actif qu'à l'état de dimère (Urnov et al., 2010). Ainsi il faut deux nucléases à doigts de zinc qui vont se lier sur l'ADN en amont et en aval de la partie que l'on souhaite déléter. Les deux domaines endonucléases de FokI vont alors se rapprocher en formant un dimère et effectuer une coupure. Celle-ci va ensuite être réparée par la ligation des deux extrémités formées par la coupure créant ainsi une délétion de quelques nucléotides qui conduira à l'inactivation du gène. Les vecteurs codant les nucléases à doigts de zinc peuvent être introduits dans les cellules ES ou directement dans le zygote.

#### **2.1.1.3.2. Knock-out conditionnel**

Les techniques décrites précédemment permettent d'inactiver ou de modifier un gène, mais ces modifications sont présentes dans toutes les cellules et dès le début du développement de l'organisme transgénique. Or certains inconvénients peuvent se présenter dans cette situation, notamment si la modification importée est létale pour la cellule ou l'embryon, ou si elle affecte de nombreux tissus et induit ainsi un phénotype complexe à analyser. Pour simplifier l'étude de ce genre de cas, il peut être intéressant de créer une souris exprimant la mutation soit à partir d'un temps déterminé, soit seulement dans certains tissus.

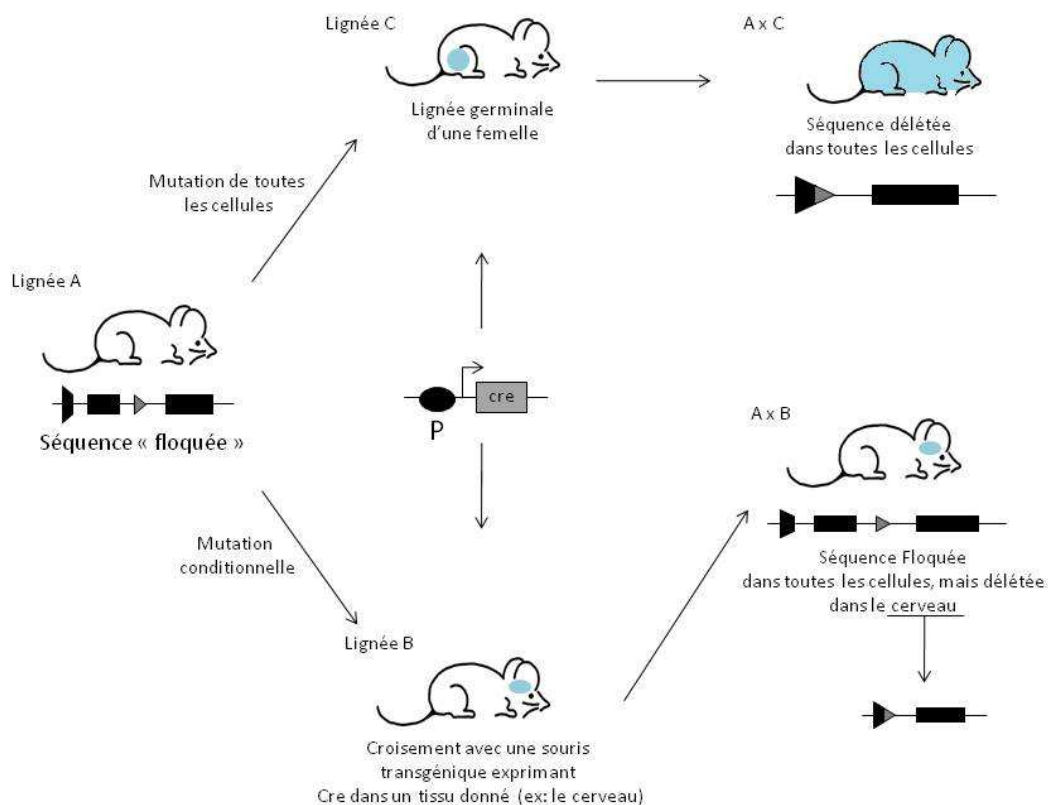
##### **2.1.1.3.2.1. Contrôle dans l'espace**

Une stratégie permettant d'exprimer une mutation seulement dans certains tissus a été développée (Gu et al. 1994). Cette stratégie est le plus souvent basée sur l'utilisation du système de recombinaison site spécifique du bactériophage P1 Cre/loxP. Premièrement, une souris transgénique portant deux sites loxP placés de façon à ce qu'ils entourent la plus grande partie du gène étudié sans perturber son activité (ceux-ci peuvent être situés par exemple dans les introns) doit être créée. Cette souris doit ensuite être croisée avec une autre souris transgénique qui exprime la recombinase Cre sous le contrôle d'un promoteur tissu-spécifique. La recombinase Cre provoque une délétion de la séquence qui se trouve entre les deux sites loxP et induit de cette manière l'inactivation du gène étudié dans les cellules où elle est exprimée (Figure 10). Il faut noter que les sites loxP ne sont pas présents naturellement chez les mammifères. Des souris exprimant Cre sous le contrôle de promoteurs variés ont été développées (Quinn et al., 2009 ; Stanford et al., 2010 ; Wang et al., 2010). Le système Flp/FRT issu de la levure agit de la même manière que Cre/loxP et est également utilisé pour déléter des séquences chez la souris (Ge et al., 2009).

##### **2.1.1.3.2.2. Contrôle dans le temps**

Il est possible de contrôler l'induction de la mutation non seulement dans l'espace, mais aussi dans le temps. Dans ce système, la protéine Cre est fusionnée avec le domaine de liaison du récepteur aux progestérones humain muté de façon à ce qu'il ne soit pas reconnu par l'hormone endogène (Kellendonk et al., 1996). Dans cette forme de fusion, la recombinase Cre n'est pas active, mais son activité peut être restaurée en présence d'un ligand

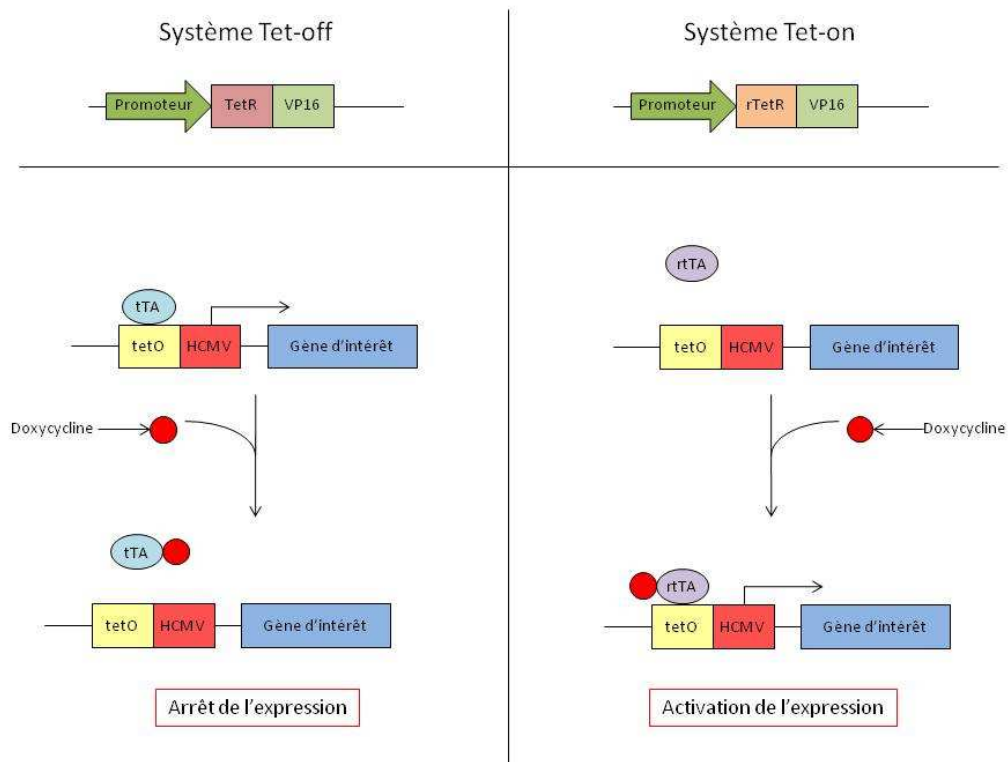
approprié (le stéroïde synthétique RU 486). Celui-ci vient se lier sur le récepteur, ce qui va entraîner un changement de conformation permettant la pénétration de Cre dans le noyau où elle peut induire la mutation ciblée. En pratique on se retrouve dans le même cas que précédemment : une souris portant deux sites loxP entourant le gène à étudier est croisée avec une souris portant le transgène de la recombinaise Cre en fusion et donc inactive. Pour l'activer il suffit d'administrer à la souris le ligand approprié dans au moment désiré. D'autres binômes récepteur/ligand peuvent être utilisés comme le récepteur aux œstrogènes muté et le tamoxifène (un analogue des œstrogènes) (Feil et al., 1996).



**Figure 10 : Mutagenèse conditionnelle basée sur le système Cre/loxP.**

Ce système est basé sur le croisement de deux lignées de souris génétiquement modifiées : la première (lignée A), obtenue après recombinaison homologe dans les cellules ES, porte dans son génome des sites loxP qui entourent le gène d'intérêt ; les lignées B et C expriment la recombinaise Cre sous le contrôle d'un promoteur donné. Si ce promoteur induit l'expression de la Cre dans la lignée germinale femelle (lignée C), le gène d'intérêt va être supprimé dans toutes les cellules des animaux issus d'un croisement entre la souris femelle C et la souris mâle A. Si le promoteur induit l'expression de la Cre dans un ou plusieurs types de cellules (lignée B), les animaux nés d'un croisement entre une souris A et une souris B seront mosaïques : le gène étudié sera supprimé seulement dans les cellules exprimant Cre. P : promoteur.

Il existe d'autres systèmes d'expression inductible, comme les systèmes Tet-Off et Tet-On qui utilisent l'opéron bactérien tétracycline (*tet*). Il s'agit de systèmes dans lesquels l'expression d'un gène d'intérêt dépend d'un activateur de transcription inductible. En pratique une lignée de souris exprimant l'activateur de la transcription sous le contrôle d'un promoteur tissu-spécifique est croisée avec une lignée où le gène d'intérêt est placé sous contrôle d'un promoteur régulé par cet activateur. Dans les deux systèmes (Tet-Off et Tet-On), l'activateur de la transcription peut être régulé à la fois de façon réversible et quantitative en exposant les animaux transgéniques à des concentrations différentes de tétracycline, ou d'un dérivé tel que la doxycycline (Kistner et al., 1996). Dans le système Tet-Off, la transcription est inactivée par la tétracycline ou la doxycycline, et à l'inverse, dans le système Tet-On, la transcription est active seulement en présence d'antibiotique (Figure 11) (Kistner et al., 1996).



**Figure 11 : Systèmes Tet-Off et Tet-On.**

a) Dans le système d'expression Tet-Off, une protéine transactivatrice (*tTA*) contrôlée par la tétracycline, régule l'expression du gène d'intérêt qui est sous le contrôle transcriptionnel du « *tetracycline responsive element* » (*TRE*). La protéine transactivatrice est composée d'une protéine bactérienne (répresseur tétracycline-*TetR*), qui reconnaît l'antibiotique tétracycline et interagit avec la séquence de l'opérateur (*tetO*), et d'un domaine transactivateur transcriptionnel VP16 (provenant du virus *Herpès simplex*). Le *TRE* est composé de l'opérateur Tet (*tetO*) fusionné à un promoteur minimal (dérivé du cytomégalo virus humain (*HCMV*)). En l'absence de tétracycline ou de doxycycline, la *tTA* se lie au *TRE* et active la



transcription du gène cible. En présence de tétracycline ou de doxycycline, la *tTA* ne peut pas se lier à la *TRE*, et le gène cible n'est pas exprimé.

b) Le système Tet-On utilise une forme mutée de la protéine transactivatrice (reverse tetracycline-controlled transactivator-*rtTA*). *rtTA* ne peut reconnaître la séquence tetO dans le TRE du transgène cible qu'en présence de tétracycline ou de doxycycline. Ainsi, dans le système Tet-On, la transcription du gène est activée par ajout d'antibiotique, et à l'inverse, l'arrêt d'administration de cet antibiotique permet d'inhiber l'expression du gène.

### **2.1.1.3.3. Inactivation à grande échelle**

Compte tenu de la puissance des études fonctionnelles par inactivation de gène chez la souris, des programmes de mutagenèse à grande échelle ont été initiés. Un premier projet basé sur la mutagenèse aléatoire par un agent chimique (l'ENU ; voir ci-dessous) (Masuya et al., 2004) a été suivi par d'autres grands projets internationaux. Les membres de l'« *International Knockout Mouse Consortium* » (IKMC) ont entrepris d'invalider tous les gènes codant des protéines chez la souris en utilisant des approches aléatoires (piégeage de gènes) ou ciblées (knock-out) dans les cellules souches embryonnaires. L'« *International Mouse Phenotyping Consortium* » (IMPC) envisage d'étudier systématiquement le phénotype des souris générées à partir des cellules souches embryonnaires développées par l'IKMC. Chaque lignée de souris va être soumise à une batterie de tests à haut débit pour l'identification de phénotypes développementaux, anatomiques, physiologiques, comportementaux et pathologiques. Quinze mille lignées de souris devraient être phénotypées d'ici 2020.

#### **2.1.1.3.3.1. Mutagenèse par l'ENU**

La N-éthyl-N-nitrosourée (ENU) est un mutagène chimique puissant utilisé chez la souris, qui induit des substitutions de bases de façon aléatoire dans le génome entier des spermatogonies (Russel et al., 1982). Une stratégie basée sur cette méthode a été mise en place pour sélectionner des souris mutantes en fonction de leur phénotype. Les souris mâles traitées à l'ENU sont croisées avec des souris femelles sauvages. La descendance est soumise à une analyse phénotypique pour sélectionner des mutations candidates. Les phénotypes dominants peuvent être détectés dès la génération F1. Pour les phénotypes récessifs il faut d'abord effectuer un croisement entre la génération F1 et des souris sauvages puis croiser leur descendance F2 avec les souris F1. Les études phénotypiques sont effectuées sur les souris provenant de ce dernier croisement (génération F3). Une fois les individus portant la mutation repérés, il faut procéder à l'identification de la mutation par clonage positionnel. Grâce à cette

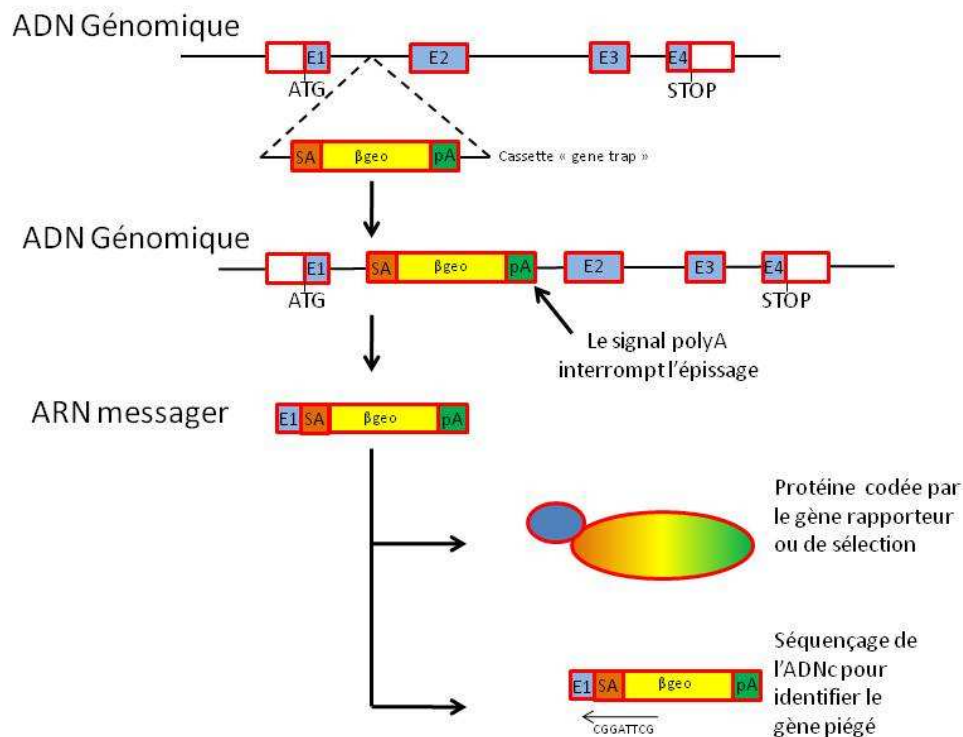
stratégie, une mutation récessive dans le gène *FBN2* (*fibrilin 2*, une composante de microfibrilles du tissu conjonctif probablement impliquée dans l'assemblage des fibres élastiques) qui se traduit par une myasthénie a été identifiée (Miller et al., 2010). Cette mutation est un modèle important pour étudier les mécanismes pathogéniques sous-jacents à la myasthénie. En outre, ce travail a permis d'identifier un phénotype, jusque-là inconnu, qui semble résulter de l'absence de protéine Fbn2.

Des banques de mutants obtenus par exposition à l'ENU ont été constituées (Masuya et al., 2011). Elles sont basées sur un système d'archivage à la fois des spermatozoïdes et de l'ADN génomique des souris mâles de la génération F1. Les archives d'ADN génomique peuvent être utilisées pour rechercher des mutations induites par l'ENU dans un gène d'intérêt. Les souris peuvent alors être produites grâce au sperme cryoconservé. Ces souris mutantes peuvent être utilisées pour étudier la fonction biologique du gène d'intérêt ou fournir des modèles de maladies humaines. Par exemple il a été montré qu'une substitution induite par l'ENU dans le gène de la N-acyltransférase 1 (*Nat1*), une enzyme qui intervient dans le métabolisme des médicaments et autres xénobiotiques, provoque une réduction significative de son activité enzymatique (Erickson et al., 2008).

#### **2.1.1.3.3.2. Approche Genetrap**

Le « *gene trapping* » (piégeage de gènes), consiste à inactiver des gènes par insertion aléatoire d'une cassette de « *gene trap* » dans le génome des cellules ES. La cassette de « *gene trap* » possède plusieurs séquences clés : un site accepteur d'épissage, un signal de polyadénylation, un gène rapporteur (habituellement le gène bactérien *lacZ*) et un gène qui permet la sélection des cellules dans lesquelles un gène a été piégé (en général il s'agit d'un gène de résistance à un facteur de sélection, par exemple le gène de la néomycine phosphotransférase) (Gossler et al., 1989). Quand la cassette s'insère dans un intron d'un gène, l'épissage a lieu entre le site donneur d'épissage situé en 3' de l'exon qui se trouve en amont de la cassette et le site accepteur d'épissage situé en 5' de la cassette (Figure 12). Le gène rapporteur sera exprimé sous le contrôle du promoteur du gène piégé. Grâce au signal de polyadénylation situé en 3' de la cassette, qui sert de signal de terminaison de la transcription, les exons du gène situés en aval de la cassette ne seront pas transcrits. Plus la cassette est insérée près de l'extrémité 5' du gène, plus la partie non transcrite du gène est grande, ce qui mènera à un produit tronqué inactif. Même si, dans la plupart des cas, l'insertion de la cassette

dans la région 5' du gène piégé mène à son inactivation complète, il arrive parfois qu'on observe une expression résiduelle du transcrit pleine longueur (Voss et al., 1998).

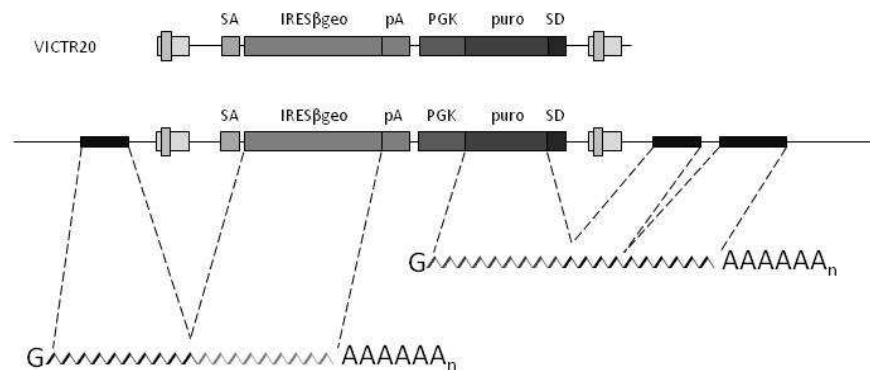


**Figure 12 : Principe de la technique « gene trap ».**

La cassette de « gene trap » s'insère de façon aléatoire dans le génome (dans le premier intron d'un gène sur cette image). La transcription du gène ainsi piégé aboutit à un ARNm tronqué. Le gène rapporteur ainsi que le gène de sélection (βgeo, fusion de lacZ et neo) sont exprimés sous le contrôle de la région régulatrice du gène piégé. Pour identifier l'intron dans lequel la cassette s'est insérée, on effectue une RT-PCR en utilisant une amorce qui s'hybride en 5' de la cassette. Le séquençage du produit obtenu permet d'identifier l'exon situé immédiatement en 5' de la cassette.

La plus grande limitation de cette méthode est que le gène de sélection n'est exprimé que si la transcription a lieu, c'est-à-dire si la cassette s'est insérée dans un gène actif. Or il peut arriver que certains gènes piégés ne soient pas actifs dans les cellules ES (par exemple le gène Btk (*Bruton's tyrosine kinase*) exprimé spécifiquement dans les lymphocytes B est actif seulement pendant la différenciation (Zambrowicz et al., 1998)). De nouveaux vecteurs de « gene trap », dans lesquels la séquence PGKpuroSD a été ajoutée, ont été développés pour contourner cette limitation (Zambrowicz et al., 1998). Cette séquence contient un promoteur PGK (promoteur du gène codant la phosphoglycérate kinase1), actif dans les lignées ES, fusionné à la séquence de la puromycine N-acyltransférase dont le signal de polyadénylation a été remplacé par un site donneur d'épissage (Figure 13). Cette amélioration permet de piéger

des gènes indépendamment de leur statut d'expression dans les cellules souches embryonnaires.



**Figure 13 : Un vecteur de « gene trap » amélioré.**

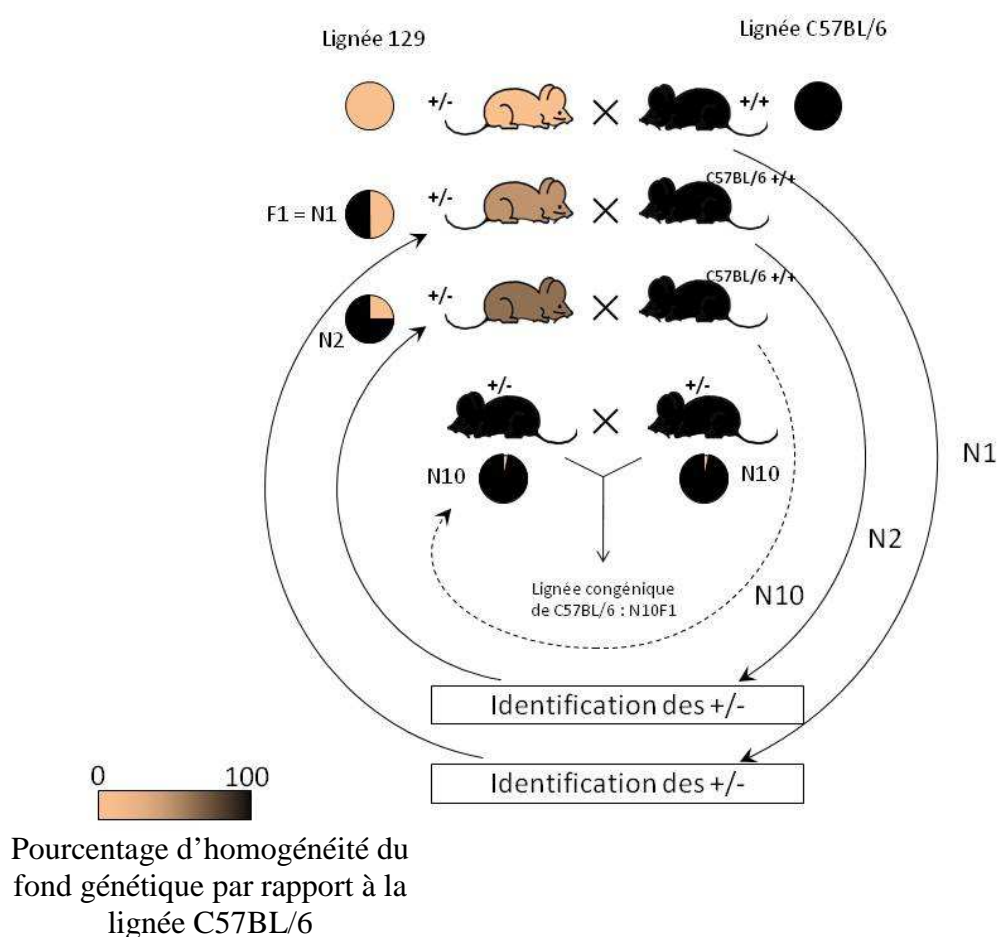
Dans ce vecteur de « *gene trap* », une cassette PGKpuro a été ajoutée en 3' de la cassette βgeo. Tandis que le gène rapporteur lacZ est exprimé sous le contrôle du promoteur du gène dans lequel le vecteur s'est inséré, la cassette PGKpuro porte son propre promoteur PGK, qui garantit l'expression dans les cellules souches embryonnaires du gène de la puromycine N-acyltransférase, ce qui permet de sélectionner toutes les cellules souches embryonnaires où un gène a été piégé (d'après Zambrowicz et al., 1998)

Des banques publiques de lignées de cellules ES murines modifiées par la technique de « *gene trap* » ont été créées par l'« *International Gene Trap Consortium* » ([www.genetrap.org](http://www.genetrap.org)). Omnibank, qui utilise un vecteur de « *gene trap* » contenant la séquence PGKpuroSD, fait partie de ce consortium (Zambrowicz et al., 1998).

Une autre approche d'inactivation de gènes à grande échelle chez la souris repose sur l'utilisation de transposons. La transposition est un processus précis par lequel un fragment d'ADN est excisé et déplacé vers un autre site dans la même molécule d'ADN ou dans un génome différent. Les transposons non-autonomes sont plus utiles comme outils de génétique parce qu'après leur insertion dans le génome ils ne peuvent pas continuer à s'exciser et à se réinsérer. Le système de transposition « *Sleeping Beauty* » a été conçu pour insérer des séquences spécifiques dans les génomes des vertébrés. Il est composé d'une transposase et un transposon (Ivics et al., 1997). Comme toutes les autres transposases de type Tc1/mariner, la transposase « *Sleeping beauty* » insère le transposon dans des régions du génome riche en AT. Pour inactiver des gènes à grande échelle chez la souris, on croise une lignée qui contient dans son génome le transposon avec une lignée qui exprime la transposase. Des événements de transposition se produisent dans les tissus de la descendance où la transposase est exprimée. Si le transposon est inséré dans un gène, il pourra conduire à son inactivation (Carlson 2003).

### 2.1.1.4. Lignées congéniques

Pour les études phénotypiques, il est recommandé d'utiliser des souches de souris bien caractérisées, notamment s'il s'agit d'une analyse de souris transgéniques. Or les cellules ES qui procurent les meilleurs rendements de transgénèse (par exemple les cellules provenant de la souche 129) sont issues de souches qui se prêtent peu à certaines études phénotypiques, en particuliers comportementales. Pour changer de fond génétique, il faut effectuer un « back-cross », c'est-à-dire des croisements successifs avec des souris sauvages ayant le fond génétique souhaité (par exemple C57BL/6). La Figure 14 schématise le processus de « congénisation » des souris transgéniques.



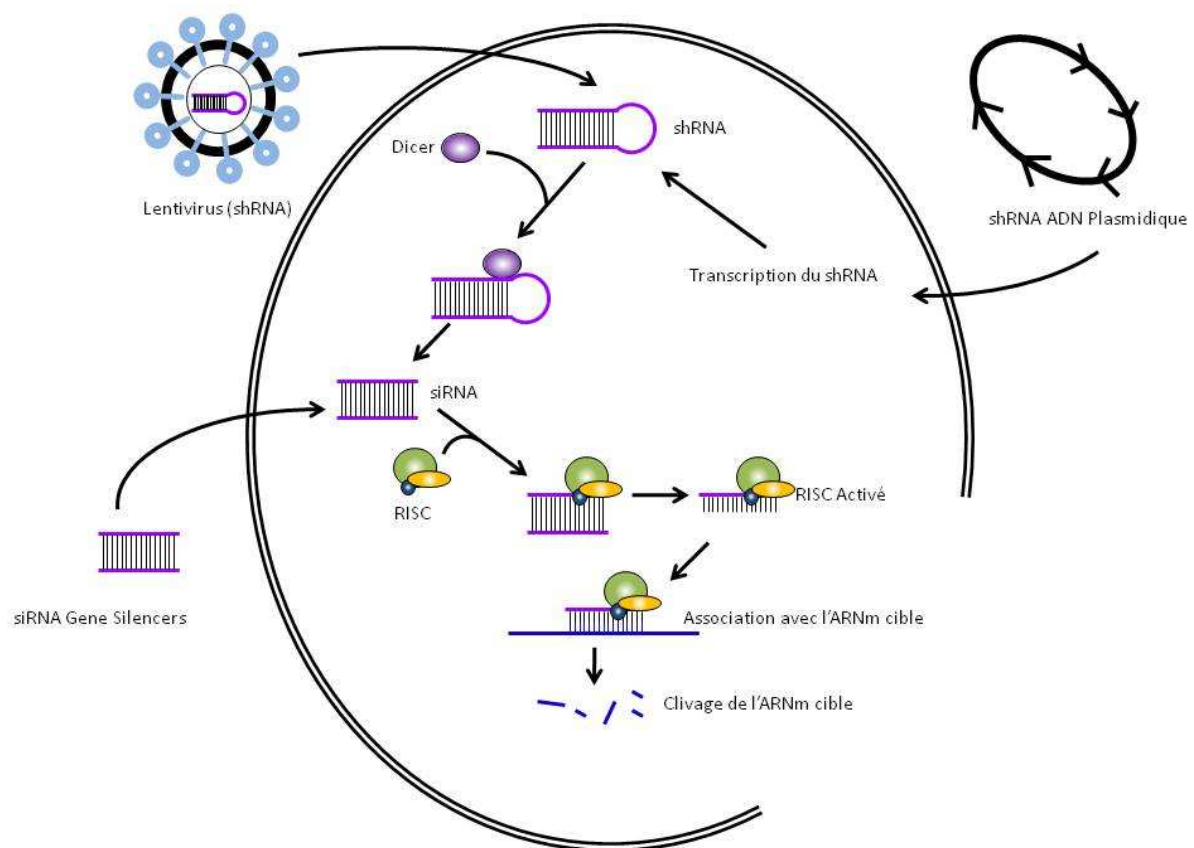
**Figure 14 : Processus de « congénisation » des souris transgéniques.**

Chaque nouvelle génération d'hétérozygotes est croisée avec des souris sauvages ayant le fond génétique souhaité. Au moins dix croisements successifs sont nécessaires à l'obtention d'une lignée congénique, c'est-à-dire à des souris avec un fond génétique considéré comme homogène.

Compte tenu du temps de génération, 2 à 3 ans sont nécessaires pour mener à terme le processus de congénisation. La stratégie dite « *speed congenic* » a été développée pour circonvenir ce problème. Dans cette méthode on analyse à chaque génération non seulement la transmission de la mutation, mais aussi des micro-satellites polymorphes entre la souche de départ et le fond génétique désiré. Les souris contenant le pourcentage le plus élevé d'allèles spécifiques du fond désiré sont sélectionnées pour le croisement suivant. En utilisant cette stratégie, seulement 5 générations (10 à 14 mois) sont nécessaires pour obtenir une lignée de souris congéniques dont le fond génétique est composé à plus de 99% du génome désiré (Wakeland et al., 1997).

### **2.1.2. Knock-down**

L'inactivation ciblée des gènes par ARN interférence (ARNi) a été développée comme une alternative au knock-out. Il s'agit d'une inhibition post-transcriptionnelle de l'expression génique. Ce mécanisme est naturellement présent chez les plantes et les animaux et il est responsable d'importantes fonctions biologiques comme la défense contre les transposons et les virus (Lindbo et al., 1993 ; Chotkowski et al., 2008). Pour diminuer le niveau d'expression d'un gène, on peut introduire dans la cellule soit directement un siRNA (petit ARN double brin) soit un shRNA (petit ARN en épingle à cheveux) ayant une séquence complémentaire à celle de l'ARN cible (Figure 15). Il faut noter que l'effet des siRNA est limité dans le temps et dans l'espace. L'ARNi via l'expression de shRNA a été appliquée avec succès dans de nombreux organismes *in vivo*, ce qui indique une grande universalité de ce mécanisme (Haley et al., 2010 ; Liu et al., 2010 ; Tinoco et al., 2010). Une perspective d'application de ce mécanisme est qu'il pourrait servir d'outil de thérapie génique (Kim et Rossi, 2007).



**Figure 15 : Vue schématique du knock-down par l'ARN interférence.**

Le processus de l'ARN interférence est déclenché par l'introduction de shRNA (*short hairpin RNA*) dans une cellule par transfection ou infection par un lentivirus. Les shRNA vont être traités dans la cellule par la ribonucléase (RNase III) nommée *Dicer*, produisant ainsi de petits ARN double-brin, d'environ 22 nucléotides, les siRNA (*small interfering RNAs*) (Paddison et al., 2002). Les siRNA s'associent avec des protéines pour former un complexe nommé RISC (*RNA-induced silencing complex*). Un seul brin du siRNA reste associé au complexe RISC. Le complexe RISC reconnaît les ARN messagers qui présentent une séquence complémentaire à ce brin et catalyse leur dégradation ; cette réaction a lieu dans le cytoplasme. Ce mécanisme peut aussi être induit en injectant dans la cellule directement de petits ARN double-brin (siRNA) qui ciblent l'ARNm endogène du gène dont on souhaite réprimer l'expression.

### 2.1.3. Surexpression

Au contraire des techniques évoquées dans les chapitres précédents, on peut étudier la fonction d'un gène en examinant l'effet de sa surexpression (ou réexpression dans le cas de gènes inactivés par knock-out ou knock-down). Pour exprimer un gène dans une cellule eucaryote, deux paramètres doivent être pris en compte : la nature du vecteur et le promoteur. Les promoteurs utilisés peuvent être d'origine virale (par exemple CMV (*cytomegalovirus immediate-early promoter*) ou SV40 (*simian virus 40 early promoter*)) ou d'origine eucaryote (par exemple EF1A (*human elongation factor 1 $\alpha$  promoter*) ou PGK (*mouse phosphoglycerate kinase 1 promoter*)). Le choix du promoteur dépend surtout du niveau d'expression désiré. Il a été montré que certains promoteurs sont forts indépendamment du type cellulaire, comme EF1A ou SV40, et que d'autres sont constamment faibles, comme PGK. En revanche, il existe des promoteurs dont la puissance dépend du couple cellule/promoteur, comme CMV (Qin et al., 2010).

Il est parfois difficile de transfecter efficacement des cellules en culture. Ces difficultés peuvent être dues à la toxicité du vecteur ou du gène surexprimé dans la cellule. Dans ce cas, on peut être amené à trouver un compromis entre la toxicité et l'efficacité de la transfection. De plus, la plupart des plasmides ne sont pas répliqués de façon épisomale chez les eucaryotes : après transfection, l'expression décline. On parle d'expression transitoire. Pour circonvenir cette difficulté, on peut chercher à établir une lignée dans laquelle le transgène sera exprimé de façon stable. Pour obtenir l'expression stable du transgène, on sélectionne une cellule ayant intégré le plasmide dans son génome (le vecteur doit posséder un marqueur de sélection eucaryote) et l'on procède à une étape de clonage.

Pour exprimer un gène dans une cellule eucaryote on peut aussi utiliser un vecteur viral (lentivirus ou adénovirus). Cette technique peut être employée *in vivo*. On choisit le virus selon son tropisme, et suivant que l'on vise ou non à intégrer le transgène dans le génome de la cellule ciblée. Leur utilisation impose des normes strictes de sécurité : on utilise des virus rendus incapables de se répliquer après infection de leur cellule cible. Les adénovirus peuvent être utilisés dans différents types cellulaires. Leur ADN ne s'intègre pas dans le génome de la cellule cible. Ils sont capables d'incorporer des transgènes de grande taille ( $\leq 8$  kb) avec un haut niveau d'expression, et sont faciles à produire. En revanche les réponses immunitaires contre l'adénovirus peuvent limiter leur emploi en thérapie génique (Nayak et Herzog, 2010). Les vecteurs lentiviraux limitent la taille du transgène à 5kb. Ils s'intègrent de façon stable dans le génome de la cellule hôte. Ils permettent l'expression de



transgènes à long terme dans des cellules non mitotiques tels que les neurones (Naldini et al., 1996).

## **2.2. Analyse spatio-temporelle de l'expression d'un gène**

Le profil spatio-temporel d'expression d'un gène peut fournir des indications sur sa fonction. Par exemple l'étude du profil spatio-temporel d'expression du gène *Neil3*, identifié comme une glycosylase de l'ADN grâce à des homologies de séquences, révèle qu'il est préférentiellement exprimé pendant le développement embryonnaire chez la souris ainsi que dans des tissus tumoraux (Hildrestrand et al., 2009), ce qui suggère un rôle dans la prolifération cellulaire. On peut établir un profil d'expression à partir d'extraits tissulaires ou directement sur coupes, par des méthodes qualitatives ou quantitatives, qui permettent de détecter le produit même du gène ou celui du transgène (ARNm, protéine ou activité enzymatique). Il est important de comparer les profils d'expression d'un gène au niveau à la fois des ARN et des protéines, en particulier s'il s'agit d'un gène exprimé dans les neurones. En effet, les ARNm sont détectés dans la région du cerveau où sont situés les corps cellulaires, mais leurs produits protéiques peuvent être localisés dans la structure de projection, comme c'est le cas pour CalDAG-GEFI (une protéine impliquée dans l'activation de la cascade des ERK/MAP kinases) qui est synthétisé par les neurones de projection du striatum et transporté vers les terminaux striatopallidaux et striatonigraux. L'immunoréactivité de CalDAG-GEFI marque l'intégralité des voies striatopallidale et striatonigree, depuis le compartiment de la matrice striatale jusqu'au pallidum et à la substance noire réticulée, tandis que ses ARNm ne sont détectés que dans le striatum (Kawasaki et al., 1998).

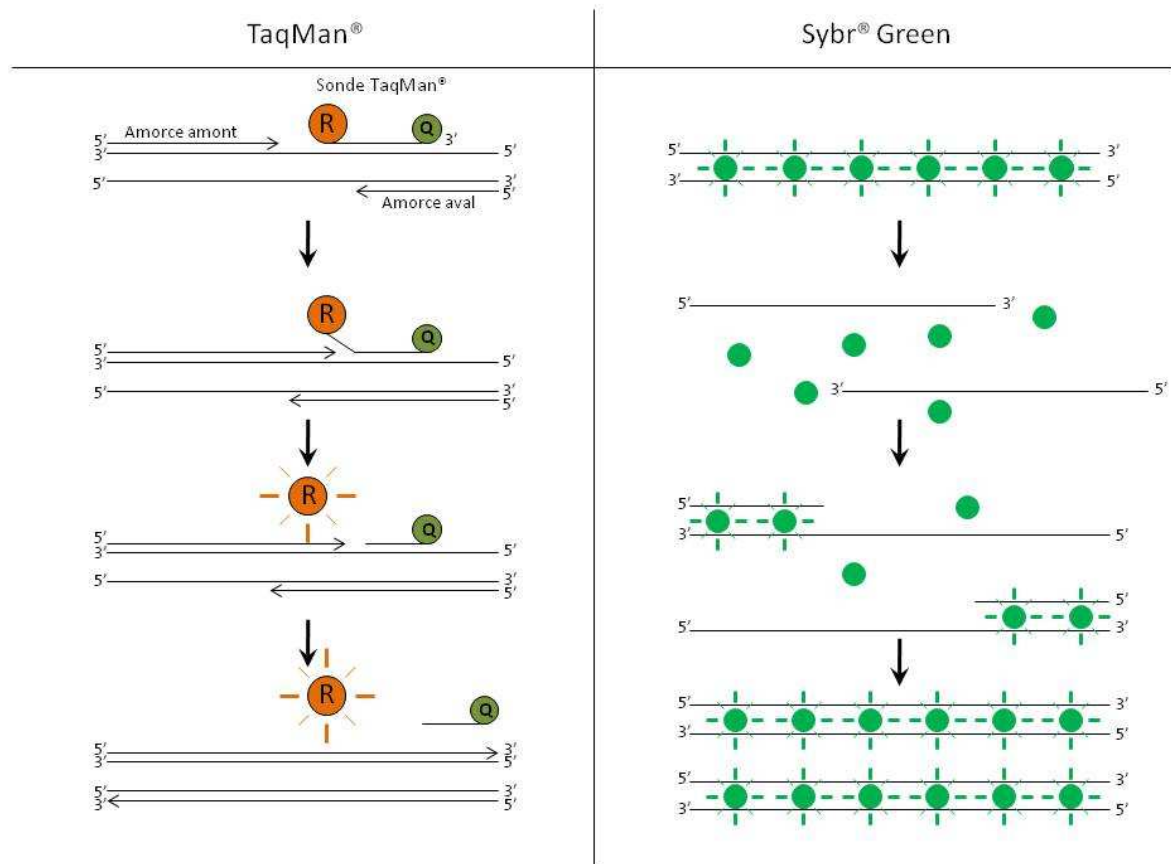
### **2.2.1. Approches moléculaires et biochimiques**

#### **2.2.1.1. A l'échelle de l'ARN**

Les ARN messagers représentent de 1 à 5% des ARN totaux extraits de la cellule. L'abondance cellulaire des ARN messagers de la plus grande partie des gènes exprimés est faible. Leur étude nécessite donc l'utilisation de techniques sensibles ainsi que la mise au point de méthodes de quantification.

Une des techniques classiques de l'analyse des ARN est le Northern blot qui consiste à séparer les ARN sur gel, à les transférer sur membrane puis à les hybrider avec une sonde complémentaire d'une partie du transcrit d'intérêt. Cette méthode fait plutôt partie des méthodes qualitatives même si l'on peut évaluer approximativement la quantité des ARN détectés en se rapportant au signal obtenu pour un ARN choisi comme témoin. Elle présente l'avantage de renseigner sur la taille des transcrits et la présence éventuelle de variants d'épissage.

Plus sensible que les autres techniques, la RT-PCR en temps réel (qPCR) est une méthode quantitative qui permet de détecter facilement même de faibles taux d'ARN. Elle comprend une étape de transcription inverse (RT) suivie d'une PCR avec des amorces spécifiques du transcrit d'intérêt. Le principe de la qPCR consiste à détecter et quantifier un rapporteur fluorescent dont l'émission est directement proportionnelle à la quantité d'amplicons générés durant la réaction de PCR, et donc à la quantité initiale de matrice. Deux techniques sont utilisées principalement pour effectuer une qPCR (Figure 16) : la technique TaqMan et la détection de l'intercalant fluorescent SYBR-Green. La possibilité de détecter toute molécule d'ADN double brin grâce au SYBR-Green peut être comprise à la fois comme un avantage et comme un inconvénient. En effet, cette technique ne requiert pas la synthèse d'une sonde spécifique. En revanche des faux positifs peuvent être produits. Au contraire, les sondes TaqMan sont spécifiques de la séquence d'ADN quantifié. Une autre application de la méthode TaqMan est la discrimination allélique. Pour cette application, des sondes spécifiques de chaque allèle sont marquées par des fluorophores différents et utilisées dans la même réaction.



**Figure 16 : Différences entre les techniques TaqMan et SYBR-Green** (D'après [www.appliedbiosystems.com](http://www.appliedbiosystems.com)).

A) Principe TaqMan. Une sonde oligonucléotidique complémentaire d'un fragment de la séquence à amplifier est construite. Elle contient un rapporteur fluorescent à son extrémité 5' et un inhibiteur de fluorescence (« *quencher* ») à son extrémité 3'. Si la sonde est intacte, la proximité de l'inhibiteur réduit considérablement la fluorescence émise par le rapporteur. Si la séquence cible est présente, la sonde s'hybride en aval de l'une des amorces et est clivée par l'activité nucléase 5' de l'ADN polymérase Taq pendant l'élongation. Le clivage de la sonde sépare le rapporteur fluorescent de l'inhibiteur de fluorescence, ce qui entraîne une augmentation de l'intensité de fluorescence proportionnelle à la quantité d'amplicons produits.

B) Principe d'utilisation du SYBR-Green. Quand le colorant fluorescent SYBR Green I est ajouté à un échantillon, il s'intercale dans toutes les molécules d'ADN double brin présents dans l'échantillon. Au cours de la PCR, l'ADN polymérase amplifie la séquence cible, et le SYBR Green I, s'intercale dans chaque nouvelle copie. A mesure que la PCR progresse, l'intensité de fluorescence augmente de façon proportionnelle à la quantité de produit de PCR formé. Pour normaliser la quantification de la fluorescence on utilise la référence passive « ROX ». Sa fluorescence passive permet de normaliser des fluctuations de fluorescence qui ne sont pas liées au signal mais qui peuvent être dues par exemple aux variations de température ou d'enregistrement au niveau de la caméra.

### **2.2.1.2. A l'échelle de la protéine**

Les protéines peuvent être analysées par Western blot, une technique traditionnelle qui consiste à séparer les protéines sur un gel de polyacrylamide, à les transférer sur une membrane support, sur laquelle les molécules recherchées sont repérées à l'aide d'anticorps spécifiques. Cependant cette technique ne permet pas une quantification très précise.

Le problème de cette technique ainsi que des autres techniques immunochimiques utilisées pour la détection des protéines est qu'elles requièrent des anticorps spécifiques. Pour tenter de le résoudre, un projet, appelé « *Human Protein Atlas* », a été initié. Il s'agit d'un programme dont le but est d'explorer de façon systématique le protéome de l'homme grâce à la production à grande échelle d'anticorps spécifiques de protéines humaines (Uhlén et al., 2005). Cet atlas contient les profils d'expression protéiques de 11260 gènes basés sur 14506 anticorps. Ce sont des anticorps de lapin produits à haut débit contre des *Protein Epitope Signature Tags* (PrESTs) spécifiques de chaque protéine et produit chez la bactérie (Agaton et al., 2003).

Une autre façon de résoudre le problème de disponibilité d'anticorps spécifiques est d'étiqueter les protéines (*tagging*). Ceci consiste à ajouter à la protéine à étudier une séquence (Tag) qui permet par exemple d'analyser sa localisation cellulaire, ou les interactions moléculaires qu'elle établit. Les tags couramment utilisés (Myc, HA, Flag) peuvent être reconnus par un anticorps monoclonal. La protéine GFP (*Green Fluorescent Protein*) est très largement utilisée pour l'étiquetage des protéines, car elle peut être facilement visualisée et permet de suivre en direct la localisation des protéines au niveau subcellulaire. Les étiquettes ajoutées à la protéine facilitent également sa purification.

## **2.2.2. Analyses histologiques**

### **2.2.2.1. A l'échelle de l'ARN**

L'hybridation *in situ* consiste à hybrider une sonde marquée spécifique du transcrit d'intérêt sur des coupes de tissu. Elle donne une information qualitative complémentaire aux approches moléculaires et biochimiques citées précédemment. On peut en effet visualiser de façon précise la présence d'un ARN donné au niveau de cellules distinctes. La technique d'hybridation *in situ* a été utilisée pour créer des atlas digitaux de l'expression génique dans le cerveau de souris, tels que l'« *Allen Brain Atlas* » (Lein et al., 2007) ou le projet « *Brain Gene Expression Map* » (BGEM) (Magdaleno et al., 2006) ou encore *Genepaint*, qui est spécialisé

dans l'étude de l'expression génique au cours du développement embryonnaire chez la souris (Visel et al., 2004).

Pour analyser à grande échelle le profil d'expression d'un gène dans différents tissus au niveau de l'ARN on peut utiliser des puces tissulaires (*tissue arrays*). Cette méthode consiste à monter sur des lames de microscope des centaines de coupes de tissus. Celles-ci vont ensuite être hybridées avec une sonde spécifique de l'ARN dont on souhaite analyser l'expression. Grâce à cette technique, le profil d'expression de gènes codant des protéines liant l'ARN a pu être établi dans différents tissus humains, et à différents âges (Masuda et al., 2009).

#### **2.2.2.2. A l'échelle de la protéine**

Si l'on souhaite localiser directement la protéine codée par le gène d'intérêt ou le transgène sur une coupe de tissu, on peut utiliser l'immunohistochimie ou l'immunofluorescence. Pour ces techniques, des anticorps spécifiques dirigés contre la protéine analysée sont indispensables. En immunohistochimie, l'anticorps secondaire est conjugué à une enzyme (par exemple la peroxydase de raifort) qui catalyse la transformation d'un substrat chromogène en produit coloré. En immunofluorescence, les anticorps secondaires sont marqués par un fluorophore. L'avantage de l'immunofluorescence est qu'elle permet d'effectuer des marquages multiples en utilisant des fluorophores différents. En revanche, elle a une sensibilité plus faible.

On peut aussi visualiser les protéines portant une étiquette GFP ou se servir de la GFP comme gène rapporteur. Le projet GENSAT (*Gene Expression Nervous System Atlas*), qui a pour but de cartographier l'expression des gènes dans le système nerveux central de la souris, utilise des lignées knock-in dans lesquelles l'EGFP et un signal de polyadénylation juste en 5' du codon ATG de chaque gène (Gong et al., 2003).

#### **2.2.2.3. Marquage Xgal**

On peut également utiliser des méthodes de détection enzymatiques si le gène d'intérêt ou le transgène a une activité enzymatique connue. C'est le cas du gène rapporteur lacZ qui code la  $\beta$ -galactosidase. Les tissus sont mis en présence du substrat Xgal, qui est hydrolysé par la  $\beta$ -galactosidase en un produit de couleur bleue. Le gène rapporteur étant placé sous le contrôle de la région régulatrice du gène d'intérêt, la présence de la coloration indique que le

promoteur du gène d'intérêt est actif. En revanche l'inconvénient majeur de cette technique est qu'elle détecte l'expression du gène rapporteur et non du gène d'origine, alors que les protéines correspondantes peuvent avoir un comportement différent (par exemple si la protéine d'origine est membranaire, sa localisation subcellulaire va être différente de celle de la  $\beta$ -galactosidase qui est soluble). Une séquence de localisation nucléaire (NLS) peut être ajoutée à la séquence du gène rapporteur pour assurer que dans le cerveau la  $\beta$ -galactosidase soit détectée au niveau de la région qui contient les corps cellulaires. Un autre problème lié à l'utilisation de lacZ comme gène rapporteur est que la  $\beta$ -galactosidase endogène produit un signal intense dans certains tissus comme l'épididyme (Turner et al., 2003).

## **2.3 Autres approches**

### **2.3.1 Analyse bioinformatique**

De nos jours une quantité très importante de données de génomique, transcriptomique et protéomique est accessible. De nombreux outils de bioinformatique ont été développés pour analyser ces informations et aider à la caractérisation fonctionnelle des gènes. Ces analyses bioinformatiques peuvent être effectuées à deux niveaux, analyses des séquences en elles-mêmes, et analyse de réseaux de gènes.

L'analyse de la séquence nucléique et protéique permet parfois d'émettre une hypothèse sur la fonction de la protéine codée par le gène d'intérêt. En effet, grâce à la séquence, on peut trouver des homologies avec d'autres gènes ou protéines dont la fonction est potentiellement connue, identifier des motifs ou domaines conservés sur la protéine étudiée et prédire sa structure.

Les recherches d'homologies consistent à comparer la séquence d'intérêt à différentes banques de séquences en utilisant un logiciel d'alignement tel que BLAST (Altschul et al., 1990). On peut effectuer un alignement entre espèces (recherche d'orthologues) ou au sein d'une même espèce (recherche de paralogues). Les principales banques de séquences utilisées sont GenBank pour les acides nucléiques et UniProt pour les protéines. GenBank est une base de données de séquences nucléiques qui contient une collection annotée de presque 200 million de séquences génomiques, d'ADNc et d'EST pour plus de 350000 organismes (Benson et al., 2011). GenBank est entretenue par le NCBI et fait partie, avec la DNA DataBank du Japon et l'EMBL *Nucleotide Sequence Database*, de l'« *International Nucleotide Sequence Database Collaboration* ». UniProt est une ressource exhaustive de séquences de protéines de haute qualité annotées manuellement ou automatiquement entre les

espèces (UniProt Consortium, 2011). UniProt est le fruit d'une collaboration entre l'« *European Bioinformatics Institute* » (EBI), le « *Swiss Institute of Bioinformatics* » (SIB) et le « *Protein Information Resource* » (PIR).

L'identification de motifs ou de domaines connus sur une protéine peuvent fournir des informations sur sa fonction (domaines de liaison, domaine catalytique s'il s'agit d'une enzyme, etc.) ou sur sa localisation dans la cellule (domaine transmembranaire, séquence d'adressage vers la mitochondrie, signal de localisation nucléaire (NLS) etc.). PROSITE est une base de données contenant environ 1600 entées (familles protéiques, domaines et sites fonctionnels). La base de données Pfam répertorie des familles de protéines (~12000), chacune représentée par des alignements multiples de séquences (Finn et al., 2010). La recherche de domaines protéiques est facilitée par des outils comme SMART (*Simple Modular Architecture Research Tool*) où plus de 500 domaines trouvés dans des protéines de signalisation, des protéines extracellulaires et des protéines associées à la chromatine sont recensés (Schultz et al., 2000).

La prédiction de la structure secondaire, voire tertiaire, de la protéine peut aussi être informative. Aujourd'hui il existe des programmes comme Phyre qui permettent de prédire la structure secondaire de façon fiable (régions avec des hélices  $\alpha$  ou des feuillets  $\beta$ , régions transmembranaires ou régions hydrophiles, qui ont des chances d'être accessibles au solvant et sont donc vraisemblablement situées à la surface de la protéine) (Kelley et Sternberg, 2009). La prédiction de structure tertiaire n'est possible que lorsqu'il existe une protéine homologue dont la conformation est connue (Wass et al., 2010).

Un réseau de régulation génétique est un système de contrôle qui régule l'expression d'un grand nombre de gènes dans un processus physiopathologique donné (Carter et al., 2010). Des outils statistiques comme GeneNetwork (Rosen et al., 2007), permettent d'analyser ces réseaux en identifiant les gènes dont l'expression varie de façon similaire et spécifique dans certains tissus, et dans certaines conditions données.

## **2.3.2. Recherche d'interactants**

### **2.3.2.1. Double hybride**

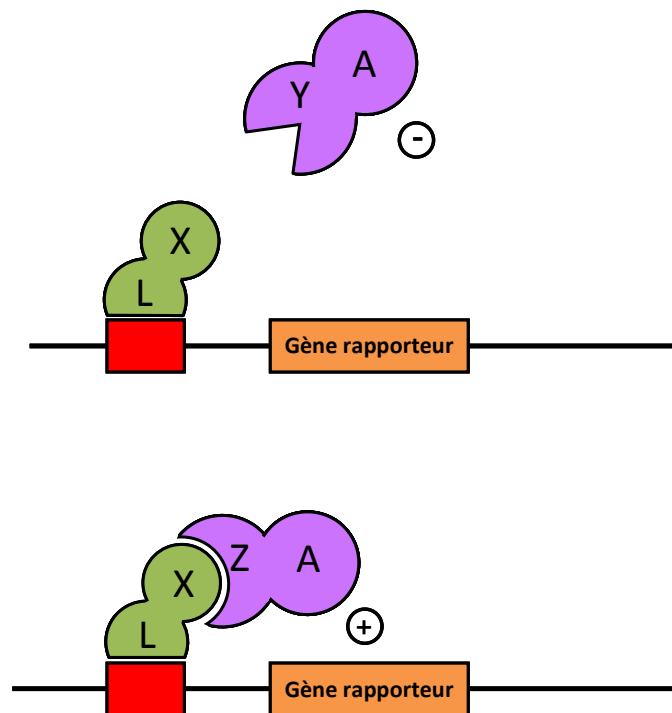
Il est souvent intéressant de connaître les partenaires d'interaction d'une protéine pour appréhender sa fonction, surtout (mais pas seulement) si la fonction de ces partenaires est connue (Jamsai et al., 2008). Le double hybride est l'une des techniques de biologie moléculaire permettant d'analyser à grande échelle les interactions physiques entre protéines

est le double hybride. Cette technique est souvent employée en raison de sa robustesse et de sa souplesse d'utilisation. Elle a été développée en 1989 par S. Fields et O. Song. Il s'agit d'un test *in vivo* dans un système cellulaire, principalement la levure *Saccharomyces cerevisiae* (Figure 17).

La méthode double-hybride présente plusieurs avantages. Tout d'abord, elle permet d'identifier aisément, à partir de colonies de levures, les protéines interagissant physiquement avec la protéine dont on cherche la fonction. De plus, la bonne conformation de la protéine peut nécessiter des modifications post-traductionnelles ou des interactions avec des molécules chaperons qui peuvent être assurées par une cellule eucaryote telle que la levure. Enfin, la sélection des protéines par la méthode double-hybride est une méthode *in vivo* contrairement à la sélection par criblage de ligands issus de la chimie combinatoire *in vitro*, souvent incapables de reconnaître leur partenaire dans un contexte intracellulaire.

En revanche un inconvénient majeur de cette technique est une grande production de faux positifs (et faux négatifs). D'après les estimations, seulement ~30% des interactions sont correctes (Deane et al., 2002). Il existe plusieurs raisons pour expliquer cette observation. Par exemple la surexpression des protéines peut mener à des interactions non-spécifiques. En outre, il peut arriver que les protéines ne subissent pas les mêmes modifications post-traductionnelles dans la levure et dans l'organisme d'origine, en particulier s'il s'agit de protéines transmembranaires. C'est pourquoi il est nécessaire de valider les interactions observées en double hybride par d'autres approches.





**Figure 17 : Système du double hybride.**

Les facteurs de transcription chez les eucaryotes possèdent dans la majorité des cas un domaine de liaison à l'ADN et un domaine d'activation de la transcription. La méthode double hybride repose sur la propriété qu'ont ces facteurs d'activer la transcription si ces deux domaines sont physiquement rapprochés, sans être nécessairement en interaction directe (par exemple s'ils sont rapprochés par un intermédiaire). Ainsi pour rechercher les interactants d'une protéine « X » deux vecteurs doivent être construits. Le premier va contenir la séquence codant la protéine « X » fusionnée à la séquence codant le domaine de liaison à l'ADN « L » du facteur de transcription GAL4. Le deuxième contient une banque d'ADNc fusionnée à la séquence codant le domaine d'activation de la transcription « A » de GAL4. Si la protéine « X » interagit physiquement avec une protéine « Z » codée par la banque, le domaine de liaison « L » et le domaine d'activation « A » vont être rapprochés, ce qui conduira à l'activation de la transcription d'un gène de sélection ou d'un gène rapporteur placé sous contrôle d'une séquence UAS (*Upstream Activating System*) spécifique du domaine de liaison de GAL4 (par exemple, un gène codant un acide aminé permettant la croissance des levures sur un milieu sélectif, ou le gène *LacZ* qui confère une coloration bleue aux levures cultivées sur un milieu contenant de l'X-gal). Après identification des colonies positives et séquençage de l'ADNc qu'elles expriment, on obtient une liste de « candidats partenaires ». Chaque partenaire doit ensuite être validé par des techniques biochimiques, afin de confirmer l'interaction moléculaire.

### 2.3.2.2. Autres techniques

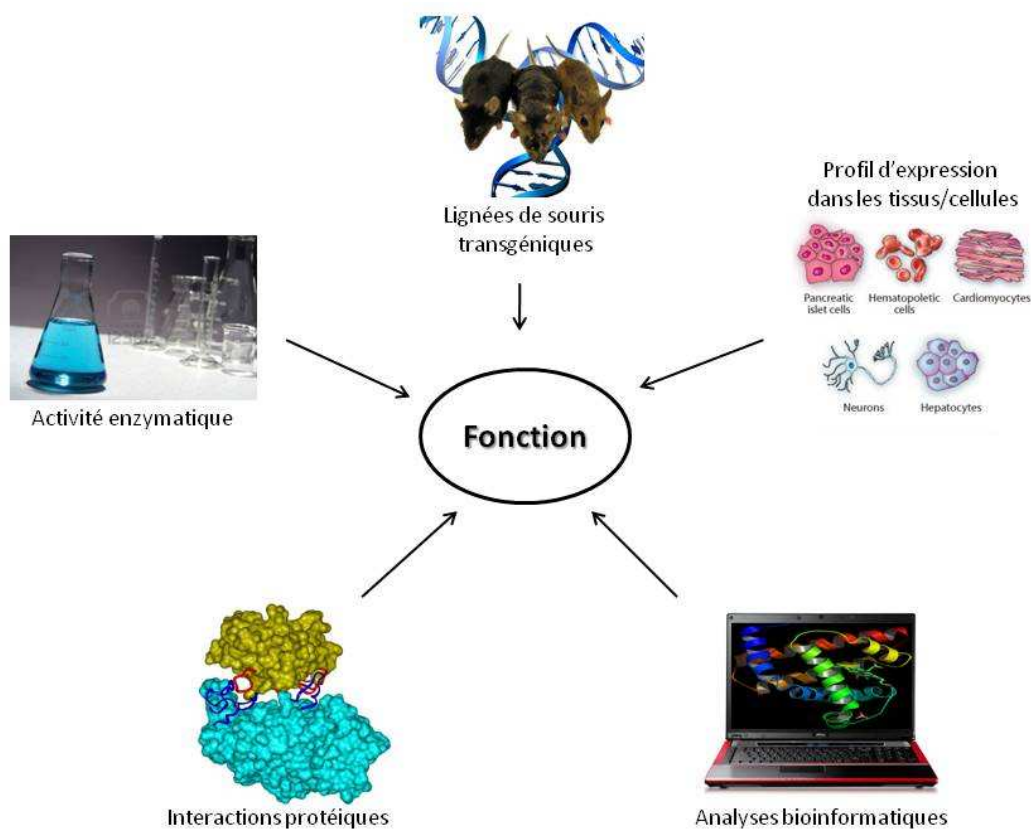
D'autres techniques ont été développées pour l'étude des interactions entre les protéines. Une de ces techniques est la co-immunoprécipitation (coIP) des protéines endogènes. Cette technique, contrairement à celle du double hybride, est difficilement réalisable à grande échelle, mais en revanche elle est d'une très grande fiabilité. Dans la pratique elle est souvent utilisée pour vérifier les interactions potentielles obtenues en double hybride. Mais elle peut aussi révéler de nouvelles interactions. Elle nécessite l'utilisation d'un anticorps spécifique dirigé contre la protéine d'intérêt. Celui-ci permet d'immunoprécipiter la protéine avec tous ses interactants liés dans un complexe. Ensuite il faut procéder à leur identification par spectrométrie de masse.

Les puces à protéines (*protein microarrays*) ont été développées comme alternative *in vitro* au double hybride. Ce sont des systèmes de liaison de ligands en phase solide où les protéines sont immobilisées sur des surfaces comme des lames de verre, des membranes, des puits de microtitration, ou des billes. Elles permettent l'analyse de centaines d'interactions en parallèle en utilisant un minimum de matériel. Grâce à leurs nombreux avantages comme la rapidité, l'automatisation, une grande sensibilité, l'économie des réactifs et la grande quantité d'informations fournies en une seule expérience, cette technique est devenue un outil central de la protéomique (Letarte et al., 2005).

De nombreuses techniques existent pour analyser les interactions entre les protéines. Des centaines de milliers d'interactions définies grâce à ces méthodes ont été répertoriées dans des bases de données spécialisées, comme DIP (*Database of Interacting Proteins*) (Xenarios et al., 2000), BioGRID (*Biological General Repository for Interaction Datasets*) (Stark et al., 2006) et STRING (*Search Tool for the Retrieval of Interacting Genes/Proteins*) (Snel et al., 2000).

### 3. Récapitulation

Comme nous avons pu le voir dans les chapitres précédents, l'étude fonctionnelle d'un gène intègre différents niveaux d'analyse : l'analyse de sa séquence, l'exploration de son profil spatio-temporel d'expression, l'observation des effets de son inactivation ou de sa surexpression, la recherche des partenaires de la protéine qu'il code (Figure 18). Dans les chapitres suivants, je vais présenter mes travaux sur la recherche de la fonction de deux marqueurs régionaux du cerveau : Agpat4 et Capucine.



**Figure 18 : Etude intégrée de la fonction d'un gène.**

### **III. Etude fonctionnelle de la Capucine**

La découverte de nouveaux biomarqueurs est souvent liée au développement d'une nouvelle technologie. Les ARNm de la Capucine (alors appelée StM20) ont été initialement identifiés par la technologie SAGE au cours d'une étude comparative des transcriptomes du cortex somatosensoriel, du striatum et du noyau accumbens chez la souris. Cette étude a montrée qu'ils étaient particulièrement abondants dans le striatum (de Chaldée et al., 2003). Une étude fonctionnelle de ce marqueur striatal a par conséquent été initiée. L'analyse de la séquence protéique de Capucine a révélé deux domaines transmembranaires putatifs à l'extrémité C terminale. Des expériences de transfection transitoire dans deux lignées de cellules de mammifères (CHO et Hela) ont montré que la Capucine fusionnée à l'EYFP ou au tag Myc est localisée dans le cis-Golgi. Un paralogue (SynDIG1) de la Capucine a également été identifié, dont le rôle dans le développement des synapses excitatrices à récepteurs AMPA a été récemment montré (Kalashnikova et al., 2010). Le profil d'expression de la Capucine a été analysé chez des souris sauvages par RT-PCR quantitative et hybridation *in situ*. Les données de RT-PCR confirment que l'expression de la Capucine est fortement enrichie dans le striatum. L'hybridation *in situ* a révélé que le gène de la Capucine est exprimé surtout dans la partie dorsolatérale du striatum (de Chaldée et al., 2006). L'enrichissement striatal des ARNm de la Capucine a également été montré chez l'homme par RT-PCR (Brochier et al., 2008). Le striatum étant la structure qui est principalement touchée au cours de la maladie de Huntington, l'expression du gène de la Capucine a été explorée dans ce contexte physiopathologique. L'abondance des ARNm de la Capucine est significativement plus faible dans plusieurs modèles murins de cette maladie que chez les témoins (de Chaldée et al., 2006 ; Desplats et al., 2006). Ces résultats indiquent que la Capucine pourrait jouer un rôle dans la pathogénèse de la maladie de Huntington.

#### **1. Le striatum**

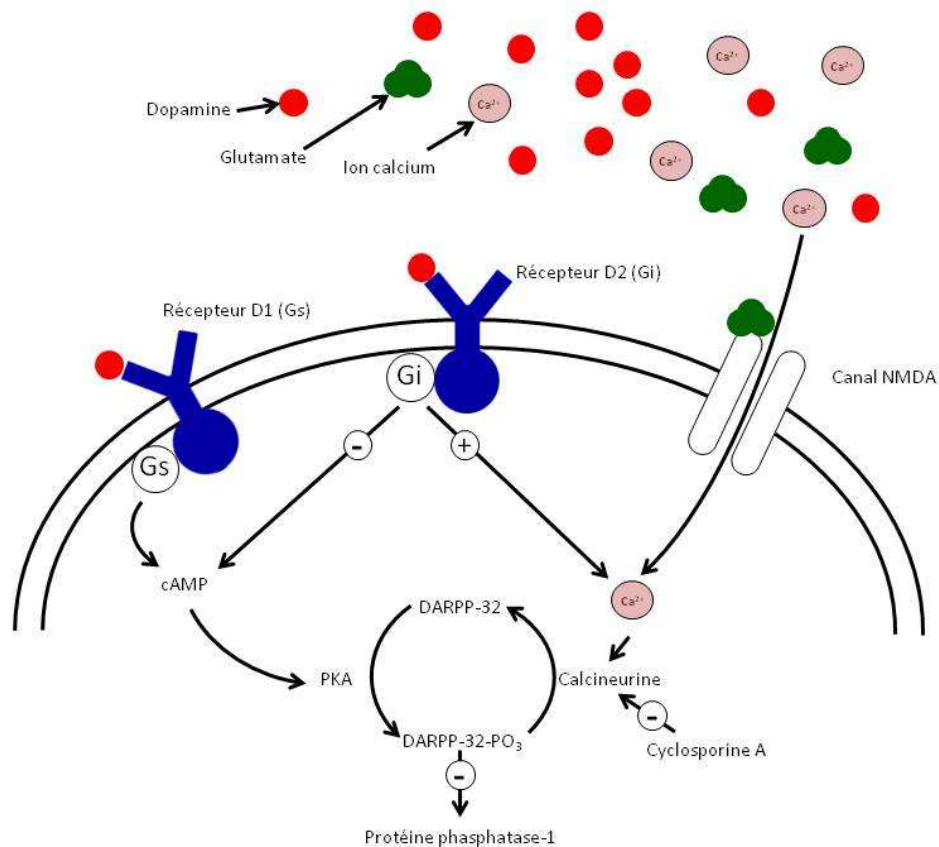
##### **1.1. Structure**

Le striatum est un noyau gris central volumineux. Chez les primates, il est composé de deux structures : le noyau caudé et le putamen, séparés l'un de l'autre par les fibres de la capsule interne. Ce passage de fibres lui confère son aspect strié, d'où son nom de striatum. Le noyau caudé et le putamen ont une origine embryologique commune et, au niveau

antérieur, ils ne sont que partiellement séparés par la capsule interne dorsalement et sont unis par le noyau accumbens au niveau ventral. Le striatum est fonctionnellement divisé en striatum dorsal et ventral. Le striatum dorsal est constitué par la plus grande partie du noyau caudé et du putamen et alors que le striatum ventral est en plus formé par le noyau accumbens et les tubercules olfactifs (Heimer et al., 1987).

Histologiquement, le striatum est constitué de deux groupes principaux de neurones : les neurones épineux de taille moyenne et les neurones non épineux, de taille moyenne ou grande. Les neurones épineux de taille moyenne représentent 90-95% des neurones striataux (Handley et al., 2006). Leur corps cellulaire fait environ 20-25µm de diamètre et leurs dendrites sont ramifiées et densément chargées d'épines (Wilson et Groves, 1980). Il s'agit de neurones de projection qui synthétisent le GABA (acide gamma-aminobutyrique), principal neurotransmetteur inhibiteur du système nerveux central. Ils sont la cible majeure des afférences extrinsèques reçues par le striatum. Les entrées corticales et thalamiques, excitatrices (glutamatergiques), forment un contact synaptique principalement avec les têtes des épines (Bouyer et al., 1984). Les fibres dopaminergiques des groupes de cellules du mésencéphale sont une autre entrée extrinsèque. Elles forment un contact synaptique surtout avec le cou des épines dendritiques et sur les axes dendritiques interépinaux (Bouyer et al., 1984). Les autres afférences que reçoivent les neurones épineux de taille moyenne proviennent principalement des interneurons du striatum (Wilson et Groves, 1980).

Les neurones épineux de taille moyenne expriment une phosphoprotéine essentielle à leur fonction : DARPP-32 (*Dopamine- and cAMP-Regulated PhosphoProtein, 32kDa*). Il s'agit d'une protéine clef de la voie de signalisation dopaminergique, qui module le flux d'informations à travers le striatum (Ouimet et al., 1984 ; Walaas et Greengard, 1984). Ce contrôle de l'état d'excitabilité des neurones épineux de taille moyenne se fait *via* deux sous-types de récepteurs couplés aux protéines G : les récepteurs dopaminergiques D1 et D2 (Figure 19) (Nishi et al., 1997). Selon leurs cibles privilégiées et leurs caractéristiques neurochimiques, les neurones épineux de taille moyenne peuvent être divisés en deux types : les neurones synthétisant l'enképhaline, exprimant le sous-type D2 des récepteurs dopaminergiques et le récepteur A2A de l'adénosine, qui projettent sur le globus pallidus externe, et ceux synthétisant la substance P et la dynorphine et exprimant le sous-type D1 des récepteurs dopaminergiques, qui cible le globus pallidus interne et la substance noire réticulée (Wang et al., 2006).



**Figure 19 : Schéma de la régulation de la phosphorylation de DARPP-32 par la dopamine.**

L'activation des récepteurs D1 (couplés à la protéine Gs – stimulatrice) augmente la concentration d'AMPc, conduisant à l'activation de la PKA (*cAMP-dependent protein kinase*) et à la phosphorylation de DARPP-32. Sous sa forme phosphorylée, DARPP-32 est un inhibiteur puissant de la protéine phosphatase-1. L'activation des récepteurs D2 (couplés à la protéine Gi – inhibitrice) diminue la phosphorylation de DARPP-32 par deux mécanismes : l'un entraîne une diminution de la concentration d'AMPc qui conduit à une diminution de l'activité de la PKA, l'autre provoque une augmentation intracellulaire de  $Ca^{2+}$  qui active la calcineurine.

Les neurones non épineux, de taille moyenne ou grande, représentent moins de 10% des neurones striataux. Ce sont des interneurons qui jouent un rôle fondamental dans l'organisation du striatum (Jain et al., 2001). Les neurones non-épineux de grande taille utilisent l'acétylcholine comme neurotransmetteur. Plusieurs types de neurones non-épineux de taille moyenne ont été décrits selon les neurotransmetteurs utilisés. Ils comprennent ceux qui contiennent la somatostatine et le neuropeptide Y et ceux qui contiennent la parvalbumine et la calrétinine (Cauli et al., 2000).

A l'échelle microstructurale, le striatum est organisé en deux compartiments : les striosomes et la matrice. La matrice représente 85% du volume striatal tandis que les striosomes en occupent 15% (Graybiel et Ragsdale, 1978). Les striosomes sont caractérisés par l'expression de récepteurs  $\mu$  aux opiacés, de la substance P et de la dynorphine (Jain et al., 2001). Ils sont entourés par la matrice riche en calbindine, somatostatine, et en acétylcholinestérase (Graybiel et Ragsdale, 1978). Cette organisation microstructurale a une importance fonctionnelle au niveau des ganglions de la base puisque l'organisation des connexions striatales suit aussi ce schéma striosomes/matrice (Gerfen, 1992). Les afférences reçues par la matrice proviennent principalement du thalamus ainsi que des régions associatives et sensori-motrices du cortex (préfrontales, temporales et pariétales) (Donoghue et Herkenham, 1986). Les striosomes sont essentiellement innervés par des fibres afférentes du cortex orbitaire et des régions limbiques (cingulaires et amygdaliennes) (Donoghue et Herkenham, 1986). Les efférences de la matrice projettent principalement sur le pallidum et la substance noire réticulée, et les efférences des striosomes projettent essentiellement sur la substance noire compacte (Graybiel et Ragsdale, 1979).

## **1.2. Fonction**

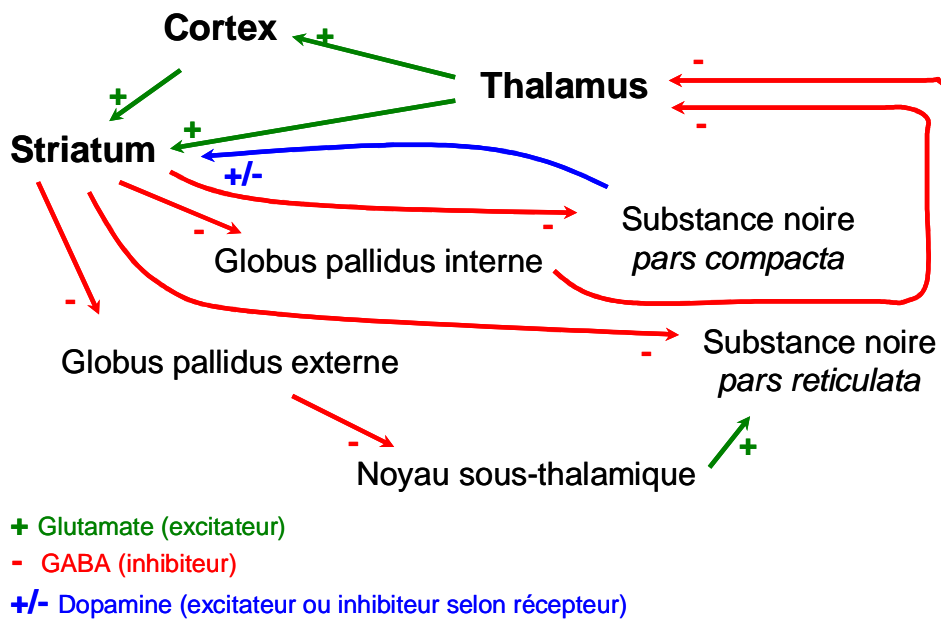
Le striatum est principalement impliqué dans la motricité volontaire mais également dans la motricité automatique. Il joue un rôle crucial dans le comportement psychomoteur, en étroite collaboration avec le cortex cérébral. Son rôle dans les fonctions motrices a été déduit des lésions neuropathologiques observées dans les maladies neurodégénératives qui affectent le contrôle des mouvements, comme la maladie de Parkinson et la maladie de Huntington. N'ayant aucun accès direct aux motoneurons spinaux, les ganglions de la base (ou noyaux gris centraux) influencent le comportement moteur en agissant sur le cortex moteur via le thalamus (motricité volontaire) et sur le tronc cérébral (motricité automatique) (Parent, 1990).

Le striatum est connu pour son rôle dans la planification et la modulation des mouvements, mais est également impliqué dans des processus cognitifs variés impliquant les fonctions exécutives. Chez l'homme le striatum est activé par des stimuli associés à une récompense, mais aussi par le caractère aversif de nouveaux stimuli, inattendus ou intenses, et par les indices associés à ces événements.

### **1.2.1. Fonction du striatum au sein des ganglions de la base**

Les ganglions de la base forment plusieurs circuits complexes auxquels participent les neurones épineux de taille moyenne. Le rôle du striatum est prépondérant au sein de ces circuits, où il constitue une sorte de "porte d'entrée". Il est en effet la cible majeure des afférences corticales aux ganglions de la base. Le principe général du circuit principal (cortico-striato-thalamo-cortical) est le suivant : l'ensemble du cortex projette sur le striatum (Divac et al., 1977), l'information est ensuite acheminée à travers les noyaux gris centraux (striatum, globus pallidus, substance noire) jusqu'au thalamus, dont les neurones prémoteurs projettent sur le cortex moteur qui assure l'exécution du mouvement (Figure 20). Deux voies, directes et indirectes, permettent de traiter l'information entre le striatum et les autres ganglions de la base. Dans la voie directe, dont la résultante est excitatrice sur le cortex moteur, le striatum projette directement sur le segment interne du globus pallidus et la substance noire *pars reticulata*. Les neurones striataux de la voie directe expriment la substance P et le récepteur dopaminergique D1. La voie indirecte, dont la résultante est inhibitrice sur le cortex moteur, fait relais dans le segment externe du globus pallidus puis dans le noyau sous-thalamique. Les neurones striataux de la voie indirecte expriment l'enképhaline et le récepteur dopaminergique D2. La régulation du circuit est assurée par deux voies : d'une part les projections dopaminergiques de la substance noire *pars compacta* modulent l'activité striatale en inhibant la voie indirecte par l'intermédiaire des récepteurs D2 et en facilitant la voie directe par l'intermédiaire des récepteurs D1 ; d'autre part le striatum reçoit des afférences glutamatergiques (excitatrices) du thalamus (van der Kooy, 1979).





**Figure 20 : Schéma du circuit cortico-striato-thalamo-cortical.**

### 1.2.2. Les différents circuits striataux des ganglions de la base

Le striatum reçoit des afférences de l'ensemble du cortex cérébral. Ainsi, sur la base de ces afférences, le striatum peut être divisé en territoires associatif, sensorimoteur et limbique. Les circuits sensorimoteurs sont les mieux connus : la « boucle motrice », décrite dans la partie précédente, joue un rôle dans la programmation et le contrôle du mouvement et la « boucle oculo-motrice » contrôle les mouvements oculaires volontaires. Les circuits associatif et limbique pourraient être impliqués dans la gestion des processus motivationnels et émotionnels associés à l'acte moteur (Parent et Hazrati, 1995) et dans des fonctions cognitives complexes (Alexander et al., 1986).

Le fonctionnement de ces circuits peut être expliqué par deux concepts : (1) la convergence et l'intégration de l'information (Kemp et Powell, 1970; Morris et al., 2003), qui comprend le fait que quasiment l'ensemble du cortex projette sur le striatum. Selon ce concept, les informations issues de domaines fonctionnels distincts sont progressivement intégrées par les ganglions de la base (Kemp et Powell, 1970), (2) la ségrégation de différents types d'information dans des circuits parallèles et indépendants (Alexander et al., 1986).

## **2. La maladie de Huntington**

Elle a été décrite pour la première fois en 1872 par George Huntington. Il s'agit d'une maladie héréditaire (autosomique dominante) fatale qui se manifeste autour de 30-40 ans et évolue sur une période de 15 à 20 ans après l'apparition des premiers symptômes. C'est une maladie dite à anticipation : elle est de plus en plus précoce et/ou sévère au fil des générations. Sa prévalence varie selon les populations. Elle est estimée entre 4 et 7 pour 100 000 en Europe de l'ouest (Hoppitt et al., 2011). La maladie de Huntington est causée par une mutation localisée dans le gène codant la protéine Huntingtine. Les mécanismes physiopathologiques induits par la protéine mutée entraînent une dégénérescence préférentielle des neurones épineux de taille moyenne du striatum. L'atrophie progressive du striatum affecte les capacités physiques et mentales des patients. Cette maladie fait partie des hyperkinésies de type choréique (ensemble de mouvements anormaux involontaires irréguliers, souvent de grande amplitude). Les patients présentent également des troubles sévères du comportement et de l'humeur ainsi que des déficits cognitifs évoluant vers la démence jusqu'à une perte totale d'autonomie.

À l'heure actuelle, de nombreux aspects de la maladie restent incompris, en particulier les mécanismes par lesquels la mutation génétique conduit à la mort neuronale et la vulnérabilité spécifique des neurones épineux de taille moyenne du striatum. Cette pathologie reste inguérissable pour le moment et les traitements disponibles ne sont que symptomatologiques.

### **2.1. Symptômes**

La symptomatologie varie selon les individus. Il existe des cas où les troubles moteurs sont peu développés, et d'autres chez lesquels les troubles cognitifs et psychiatriques apparaissent des années avant que la maladie ne soit diagnostiquée. À l'inverse, certains patients ont de graves problèmes de motricité, alors que leurs capacités cognitives sont relativement bien conservées. La diversité et la variabilité interindividuelle des symptômes rendent difficile le diagnostic et l'estimation du début de la maladie. Généralement on considère comme début de la maladie l'apparition des premiers troubles moteurs, qui se

manifestent classiquement par des mouvements involontaires de la face, des doigts, des pieds et du thorax (Folstein et al., 1986).

Les symptômes de la maladie de Huntington peuvent être classés en trois catégories : moteurs (chorée, dystonie, bradykinésie, troubles de la marche et de l'équilibre), cognitifs (troubles de la planification, de l'attention et troubles mnésiques), et psychiatriques (anxiété, dépression, irritabilité, agressivité, désinhibition).

Les troubles moteurs se caractérisent par une hyperkinésie involontaire (chorée) qui peut ressembler au début à des « tics de nervosité » des doigts et du visage. Au fur et à mesure de l'évolution de la maladie, de plus en plus de muscles sont engagés et ces mouvements deviennent amples, désordonnés, perturbant la motricité volontaire. Aux mouvements choréiques s'ajoutent des troubles globaux de la motricité notamment des troubles de la coordination et de l'alternance des mouvements rapides des extrémités, des troubles de l'équilibre et de la marche correspondant à un syndrome cérébelleux, parfois des troubles du tonus musculaire de type hypertonie plastique, des troubles de la respiration et de la déglutition. Les fonctions autonomes ne sont pas préservées et l'on constate fréquemment des troubles du sommeil et un contrôle défectueux des sphincters (Bruyn et Went, 1986). Des troubles oculomoteurs, notamment des saccades oculaires, sont souvent observés dans les stades précoces (Berardelli et al., 1999). De part sa nature précoce, ce symptôme constitue un bon indicateur du début et de l'évolution de la maladie (Hicks et al., 2008). L'incidence des troubles moteurs est évaluée par la section motrice de l'échelle UHDRS (*Unified Huntington Disease Rating Scale*) qui comporte trente et un critères tels que la poursuite oculaire, la rigidité des bras, la démarche, la chorée de la face. Un score compris entre 0 et 4 est attribué à chacun de ces critères. A un stade plus avancé de la maladie, la chorée laisse souvent place progressivement à d'autres troubles moteurs tels que la bradykinésie, la dystonie ou la rigidité.

En outre, les patients présentent systématiquement des troubles cognitifs ou neuropsychologiques (Caine et Fisher, 1985). On observe globalement un ralentissement psychomoteur (Panegyres, 2004) et des pertes de mémoire, des troubles de l'organisation et de la planification qui seraient liés à une incapacité croissante à coordonner et à initier des actions complexes (Caine et Fisher, 1985). La mémoire des patients atteints de la maladie de Huntington est touchée de manière égale sur toutes les périodes du passé (Brandt, 1991). Une spécificité de la maladie de Huntington, comparativement à d'autres démences amnésiques (comme la maladie d'Alzheimer), est la capacité d'utiliser des indices mnémotechniques pour faciliter la mémorisation (Brandt, 1991). Les tâches nécessitant un niveau d'attention élevé, le

passage d'une tâche à une autre, ou d'une règle à une autre, sont des exercices moins bien réussis par les patients que par les sujets témoins (Hicks et al., 2008). A un stade précoce de la maladie, les patients sont peu précis dans les tests qui mettent en œuvre la planification spatiale ; par contre la prise de décision reste intacte (Watkins et al., 2000). Les patients ont également un déficit visuo-spatial dans les tâches impliquant une composante motrice mais aussi dans les tâches n'impliquant que des performances perceptives (le *Mosaic Comparison Test* qui consiste à détecter de petites différences entre deux patterns visuels) (Brandt, 1991). Les pertes cognitives s'accumulent inévitablement jusqu'à une profonde démence aux stades tardifs de la maladie.

Des troubles psychiatriques sont également présents chez les personnes atteintes de la maladie de Huntington. On observe une détérioration intellectuelle accompagnée de dépression, rarement de schizophrénie. Les personnes deviennent anxieuses, irritables, obstinées, apathiques, introverties, inattentives, présentent un émoussement affectif, ont un sentiment d'infériorité et d'échec (Bruyn et Went, 1986; Caine et Fisher, 1985). Toutefois, certains symptômes de la dépression peuvent être associés à l'incertitude d'être porteur du gène, ou à l'attente des premiers symptômes. La dépression ne peut donc pas être vue comme un indicateur significatif de la maladie (Shiwach et Norbury, 1994). Cet état dépressif peut conduire au suicide qui est 5 à 10 fois plus fréquent chez les malades que dans la population générale (Di Maio et al., 1993 ; Robins Wahlin et al., 2000).

## **2.2. Neuropathologie**

D'un point de vue neuropathologique, la maladie de Huntington est caractérisée par une dégénérescence neuronale. La région la plus sévèrement touchée est le striatum. L'atrophie du striatum résulte de la dégénérescence sélective des neurones épineux de taille moyenne. Cette dégénérescence entraîne une dilatation des ventricules, dont le volume peut atteindre 2 à 10 fois la normale (de la Monte et al., 1988). Elle s'accompagne d'une gliose astrocytaire (Vonsattel et al., 1985). En plus de la neurodégénérescence du striatum, on observe une atteinte d'autres régions cérébrales, notamment la totalité de la substance blanche, le cortex cérébral, l'amygdale, l'hippocampe, le tronc cérébral et le cervelet (Rosas et al., 2003). Cette atrophie corticale et sous corticale entraîne une perte de 25 à 30 % du poids cérébral dans les cas les plus avancés (Margolis et Ross, 2001).

Les mouvements hyperkinétiques observés dans la maladie de Huntington sont la conséquence d'une désinhibition des neurones thalamocorticaux qui excitent les neurones du cortex moteur, ce qui engendre une activité motrice excessive et incohérente (Albin et al., 1989; DeLong, 1990 ; Parent, 2006). Cette désinhibition résulte de la dégénérescence des neurones striataux à l'origine de la voie indirecte. Ces neurones sont en effet affectés plus tôt et plus sévèrement que ceux de la voie directe (Albin et al., 1989). Ceci pourrait expliquer le fait que les patients atteints de la chorée de Huntington sont d'abord hyperkinétiques (choréiques) et deviennent progressivement akinétiques et rigides dans les stades plus avancés (Mahant et al., 2003).

Le dysfonctionnement des circuits striataux est à l'origine des troubles moteurs mais aussi des troubles cognitifs dont souffrent les personnes atteintes de la maladie de Huntington. Les troubles cognitifs, qui touchent en particulier les fonctions exécutives (Snowden et al., 2003), seraient en effet liés au dysfonctionnement des circuits associatif et limbique (Alexander et al., 1986). Autant les mécanismes physiologiques rendant compte de l'hyperkinésie typique de la maladie de Huntington sont assez compatibles avec le modèle de fonctionnement des ganglions de la base présenté plus haut, autant les processus sous-tendant les déficits cognitifs s'expliquent difficilement par la seule désinhibition thalamo-corticale consécutive à la perte des neurones de projection du striatum.

### **2.3. Génétique**

La maladie de Huntington présente un mode d'hérédité de type autosomique dominant. Le locus morbide a d'abord été cartographié sur l'extrémité du bras court du chromosome 4 en 1983 (Gusella et al., 1983). Le gène et la mutation responsable de cette maladie n'ont été identifiés qu'en 1993 par clonage positionnel (The Huntington's Disease Collaborative Research Group, 1993). La mutation consiste en une répétition de 36 à plus de 120 fois du trinuéclotide CAG dans l'exon 1 du gène IT15 (renommé Huntingtine) situé en 4p16.3. Cette répétition se traduit au niveau de la protéine par une région polyglutamine anormalement longue dans la partie N-terminale de la Huntingtine (Htt). Dans la population saine, le nombre de répétitions CAG est inférieur à 26 (Sadri-Vakili et Cha, 2006). La maladie a une pénétrance complète chez les individus ayant un nombre de répétitions supérieur à 37. Les allèles de 27 à 35 répétitions sont considérés comme normaux, mais ont une grande

mutabilité alors que les répétitions de 36 à 37 trinuécléotides donnent lieu à une pénétrance réduite (The American College of Medical Genetics/American Society of Human Genetics Huntington Disease Genetic Testing Working Group, 1998). L'homozygotie pour la mutation ne modifie pas l'âge de début de la maladie (Wexler et al., 1987), mais peut accélérer sa progression (Squitieri et al., 2003).

Dans les familles d'atteints, le nombre de répétitions augmente d'une génération à l'autre, surtout lors d'une transmission paternelle (Potter et al., 2004). Ce phénomène d'expansion est le support moléculaire de l'anticipation clinique. Il s'agit du premier exemple décrit de mutation dynamique. Depuis, plus de cinquante maladies à expansion de triplets ont été décrites, la plupart neurodégénératives ou neuropsychiatriques. Cependant il s'agit d'expansions de différents types de triplets qui peuvent se rencontrer dans différentes parties des gènes (codante, 5'UTR, 3'UTR, introns) et qui présentent des seuils pathologiques différents (Timchenko et Caskey, 1996).

## **2.4. Physiopathologie**

La Huntingtine est une protéine ubiquitaire de 350 kDa, qui ne possède pas d'homologie avec d'autres protéines connues. Elle est exprimée dans tous les tissus. Son niveau d'expression est le plus élevé dans le cerveau et les testicules (Borrell-Pages et al., 2006). Dans le cerveau, elle est exprimée par les neurones ainsi que les cellules gliales (Shin et al., 2005). L'invalidation du gène de la Huntingtine par knock-out est létale chez la souris (Zeitlin et al., 1995). Même si plusieurs hypothèses ont été avancées, son rôle physiologique n'est pas encore connu. Néanmoins, les nombreux travaux réalisés au cours de ces dernières années ont permis d'élucider certaines fonctions de la Huntingtine normale et mutée, dans le développement et dans la physiologie neuronale (Zuccato et al., 2010), notamment dans la régulation de la transcription (Benn et al., 2008), la dynamique de la membrane (Kegel et al., 2009), l'efficacité des mitochondries (Aoyama et al., 2006), l'exocytose, l'autophagie et le recyclage endosomal (Kegel et al., 2000). Il est très probable qu'il s'agisse d'une protéine multifonctionnelle (Harjes et Wanker, 2003). Ainsi la mutation du gène codant la Huntingtine entraîne diverses anomalies physiologiques, en particulier un dysfonctionnement mitochondrial, notamment un déficit au niveau des complexes I et II/III de la chaîne respiratoire (Browne et al., 1997) ; des anomalies du transport des neurotransmetteurs et du BDNF (*Brain-Derived Neurotrophic Factor*), facteur normalement sécrété dans le striatum

par les neurones corticaux et qui sert de support trophique aux neurones striataux (Massouh, 2007) ; une activation anormale de la voie JNK (pour c-Jun N-terminal Kinase) (Liu, 1998). Cette voie est impliquée dans les processus apoptotiques et son activation est spécifique au striatum (Garcia et al., 2002). La dopamine libérée dans le striatum pourrait aussi jouer un rôle dans la vulnérabilité spécifique des neurones de cette structure (Charvin et al., 2005). Certaines de ces anomalies donnent donc des indices sur les raisons pour lesquelles le striatum est spécifiquement atteint.

La Huntingtine mutée est caractérisée par un domaine riche en glutamines au niveau de son extrémité N-terminale. Les mécanismes par lesquels ce domaine polyglutamine rend la protéine toxique et cause la dégénérescence préférentielle des neurones majoritaires du striatum n'est toujours pas claire (Charvin, 2005). En outre les symptômes de la maladie de Huntington ne doivent pas être uniquement considérés comme la conséquence d'une perte neuronale, mais également comme la conséquence d'un dysfonctionnement cellulaire qui reste longtemps asymptomatique (Sugars et Rubinsztein, 2003). Rangone et al. (2004) proposent que le rôle de la protéine mutée dans l'apparition de la maladie se situe à deux niveaux. D'une part, la protéine mutée serait toxique en elle-même et d'autre part, la mutation entraînerait une perte de fonction.

Des observations réalisées sur des cerveaux de patients *post-mortem* ont mis en évidence la présence d'agrégats cytoplasmiques et d'inclusions nucléaires dans les neurones atteints. Ces agrégats/inclusions sont essentiellement constitués de fragments N-terminaux de la Huntingtine mutée contenant l'expansion anormale de polyglutamine (Rangone et al., 2004). On retrouve aussi des agrégats/inclusions dans les cellules gliales (Shin et al., 2005). Selon plusieurs auteurs, ces agrégats/inclusions seraient toxiques en tant que tels et conduiraient à la mort de la cellule (Charvin, 2005; Steffan et al., 2000). Pour d'autres, il pourrait s'agir d'une stratégie neuroprotectrice : la forme soluble de la Huntingtine mutée serait toxique et son agrégation protégerait les neurones (Borrell-Pages et al., 2006; Saudou et al., 1998). Cependant il existe une troisième hypothèse d'après laquelle les agrégats/inclusions ne joueraient pas de rôle (ni toxique ni protecteur) dans la dégénérescence des neurones (Klement et al., 1998).

## **2.5. Traitement de la maladie**

A ce jour, il n'existe aucun traitement curatif ou permettant d'arrêter l'évolution de la maladie de Huntington. Cependant trois voies thérapeutiques complémentaires sont

actuellement envisagées. La première est le traitement physiopathologique médicamenteux de la neurodégénérescence. La deuxième est la thérapie génique neuroprotectrice. La dernière voie est la thérapie cellulaire substitutive qui vise à remplacer les neurones détruits par des neurones sains. Les traitements symptomatiques sont peu nombreux et peu efficaces ; aucun traitement connu ne retarde le déclenchement de la maladie ni ne ralentit son évolution (Hersch et Rosas, 2008). Les mouvements hyperkinétiques sont atténués par des antagonistes des récepteurs dopaminergiques, et les troubles psychiatriques par des neuroleptiques, des anxiolytiques et des antidépresseurs. Cependant, ces traitements pharmacologiques n'ont aucune influence sur les symptômes cognitifs ni sur la progression de la maladie.

Une des principales approches de la thérapie génique est la répression de l'expression génique par l'ARNi. Le « silencing » spécifique de l'allèle muté de la Huntingtine par ciblage de SNP associés (van Bilsen et al., 2008 ; Lombardi et al., 2009 ; Pfister et al., 2009) et la réduction partielle de l'expression à la fois de la Huntingtine sauvage et mutante (Boudreau et al., 2009 ; Drouet et al., 2009) représentent deux approches thérapeutiques potentielles. Des expériences de *silencing* allèle-spécifique conduisant à une diminution sélective de l'expression de la Huntingtine mutée de 50 à 80% ont démontré un bénéfice significatif sur la neuropathologie et le comportement moteur dans des modèles rongeurs (Boudreau et al., 2009 ; Drouet et al., 2009 ; Harper et al., 2005 ; Rodriguez-Lebron et al., 2005 ; Machida et al., 2006 ; DiFiglia et al., 2007). Le ciblage de la Huntingtine sauvage et mutante (par des shRNA et des siRNA) a été évalué *in vivo*. La suppression de l'ARNm de la Huntingtine sauvage dans le striatum de souris adulte est tolérée pendant au moins 4 mois dans 70% de cas (McBride et al., 2008). La réduction partielle de l'expression de la Huntingtine mutée dans des modèles rongeurs diminue la neuropathologie, améliore le comportement moteur et la survie (Boudreau et al., 2009 ; Drouet et al., 2009 ; Harper et al., 2005 ; Rodriguez-Lebron et al., 2005 ; Machida et al., 2006 ; DiFiglia et al., 2007).

La thérapie cellulaire peut être considérée comme une innovation clinique majeure. Elle consiste à injecter de cellules vivantes (greffe de tissu striatal fœtal) ou à administrer des facteurs neurotrophiques dans le but de ralentir la progression de la maladie en remplaçant les neurones détruits. La démonstration d'un effet fonctionnel de greffes intracérébrales de tissu neural a été apportée pour la première fois en 1979 par deux équipes dans le cadre de la maladie de Parkinson (Bjorklund et Stenevi, 1979 ; Perlow et al., 1979). Cependant, bien que les greffons semblent se développer normalement, les premiers essais cliniques ne montrent pas d'amélioration cognitive ou motrice satisfaisante (Bachoud-Lévi et al., 2006 ; Kim et al., 2008).



## **2.6. Modèles de la maladie de Huntington**

Il existe différents modèles de la maladie de Huntington. Parmi les plus couramment utilisés on peut citer des modèles pharmacologiques (par exemple l'administration d'acide 3-nitropropionique chez les rongeurs, et les primates non humains ; il s'agit d'un inhibiteur irréversible de la succinate déhydrogénase (complexe II mitochondrial) entraînant une vulnérabilité des neurones striataux (Palfi et al., 1996)) et des modèles génétiques. Dans cette partie on évoquera uniquement les modèles génétiques.

### **2.6.1. Souris transgéniques**

L'invalidation du gène de la Huntingtine par knock-out étant létale chez la souris (Zeitlin et al., 1995), il a été suggéré que la mutation à l'origine de la maladie de Huntington entraîne un gain de fonction plutôt qu'une perte. Dès lors, l'introduction d'une copie du gène codant la Huntingtine humaine muté dans le génome murin devrait être suffisante pour produire un phénotype. Les lignées de souris transgéniques R6 ont été créées en 1996 (Mangiarini et al., 1996). Ces souris ont intégré dans leur génome les séquences du promoteur et de l'exon 1 de la Huntingtine humaine portant une répétition d'environ 115 à 150 triplets CAG. Les souris de la lignée R6/2 sont les plus étudiées et les mieux caractérisées. Elles présentent un phénotype neurologique progressif qui apparaît vers l'âge de trois semaines avec de nombreuses caractéristiques de la maladie de Huntington, notamment des mouvements choréiques, des mouvements involontaires stéréotypés, des tremblements et des crises d'épilepsie, ainsi que des troubles non-moteurs tels qu'un déficit d'apprentissage qui augmente avec l'âge, ainsi qu'une baisse progressive du poids (Mangiarini et al., 1996). Les animaux meurent prématurément (à l'âge d'environ 13 semaines), mais ne montrent pas de dégénérescence neuronale massive (Mangiarini et al., 1996). Des inclusions intranucléaires et des agrégats cytoplasmiques ont également été observés, non seulement dans les neurones mais aussi dans d'autres organes tels que les muscles squelettiques, le cœur et le foie. Ils peuvent servir de marqueur précoce de la maladie, car leur formation précède l'apparition des déficits comportementaux dans les lignées R6 (Davies et al., 1997).

Des lignées transgéniques contenant l'ADNc complet du gène de la Huntingtine humaine mutée ont été développées (par exemple Reddy et al., 1998). Dans cette étude, les souris avec 48 ou 89 répétitions du triplet CAG manifestent un dysfonctionnement

comportemental et moteur progressif avec une perte neuronale et une gliose dans le striatum, le cortex cérébral, le thalamus et l'hippocampe.

Des lignées de souris knock-in ont également été produites en insérant des répétitions du codon CAG dans le gène murin sauvage. Ces souris ne montrent des déficits comportementaux que très tardivement (Menalled et al., 2002 ; Wheeler et al., 2002) et des signes de mort neuronale n'ont été observés que dans l'une de ces lignées (Wheeler et al., 2002). Des inclusions intranucléaires ont également été observées. Ces souris sont plus rarement utilisées que les souris R6/2 notamment parce que leur phénotype est moins sévère et apparaît tardivement. Cependant ce développement progressif du phénotype en fait peut être un meilleur modèle de la maladie de Huntington.

### **2.6.2. Les modèles lentiviraux**

Une stratégie basée sur l'administration intrastriatale de vecteurs lentiviraux codant un fragment N-terminal de la Huntingtine humaine mutée a été utilisée pour modéliser la maladie de Huntington chez le rat (de Almeida et al., 2002) et la souris (Faideau et al., 2010). Une pathologie progressive avec mort cellulaire a été observée au cours des six mois après l'infection. La pathologie apparaît d'autant plus tôt et est d'autant plus sévère que les fragments de Huntingtine sont plus courts, que les expansions CAG sont plus longues et que les niveaux d'expression sont plus élevés. Ceci pourrait expliquer le fait que la dégénérescence des neurones soit observée beaucoup plus tôt dans les modèles viraux que chez les souris transgéniques. En effet le niveau d'expression de la Huntingtine mutée est plus élevé dans les modèles lentiviraux (Saudou et al., 1998). La progression des signes pathologiques a été examinée dans ce modèle (Zala et al., 2005). Un mois après l'infection des agrégats neuritiques apparaissent puis des inclusions intranucléaires et des anomalies morphologiques des neurites. Deux mois après l'infection une réelle perte neuronale a été confirmée dans le striatum (Zala et al., 2005). Un des inconvénients de ce modèle est qu'il s'agit d'un effet aigu et limité dans l'espace.

### 3. Article N° 1

Une lignée de souris knock-out pour le gène de la Capucine a été construite au laboratoire peu avant mon arrivée en thèse. Les souris de cette lignée sont apparemment en bonne santé, fertiles, ne présentent pas de mortalité précoce ni d'anomalie comportementale évidente. Au cours de ma thèse, j'ai participé à la mise en place d'une stratégie de génotypage de ces souris par PCR multiplex. J'ai également analysé le profil spatio-temporel d'expression de la Capucine grâce au gène rapporteur codant la  $\beta$ -galactosidase chez les souris knock-out. J'ai aussi examiné si l'absence de la Capucine chez les souris knock-out était compensée par la surexpression de son paralogue SynDIG1. J'ai enfin initié une production d'anticorps par la méthode de *phage display*. Cette dernière partie ne figure pas dans la publication et va être évoquée dans un chapitre indépendant.

Capucine est un marqueur spécifique du striatum, et son expression est hyporégulée dans des modèles de la maladie de Huntington. Par conséquent, nous avons examiné son rôle potentiel dans le développement de cette maladie au cours d'une collaboration. Nous avons analysé l'effet de perturbations de l'expression de la Capucine chez la souris sur la neurotoxicité induite par l'injection dans le striatum d'un vecteur lentiviral codant un court fragment N-terminal de la Huntingtine mutée. L'expression de la Capucine a été perturbée par différentes approches : déficit constitutif dans la lignée knock-out que nous avons générée ; ré-expression chez les souris knock-out par injection intrastriatale d'un lentivirus exprimant la Capucine ; surexpression chez des souris sauvages par la même approche lentivirale.

## Capucin does not modify the toxicity of a mutant Huntingtin fragment *in vivo*

Laurie Galvan<sup>a,b</sup>, Nad'a Lepejová<sup>c</sup>, Marie-Claude Gaillard<sup>a,b</sup>, Carole Malgorn<sup>a,b</sup>, Martine Guillermier<sup>a,b</sup>, Diane Houitte<sup>a,b</sup>, Gilles Bonvento<sup>a,b</sup>, Fanny Petit<sup>a,b</sup>, Noëlle Dufour<sup>a,b</sup>, Patrick Héry<sup>c</sup>, Matthieu Gérard<sup>c</sup>, Jean-Marc Elalouf<sup>c</sup>, Nicole Déglon<sup>a,b</sup>, Emmanuel Brouillet<sup>a,b,\*</sup>,  
Michel de Chaldée<sup>c,\*\*,1</sup>

<sup>a</sup> CEA, DSV, I2BM, Molecular Imaging Research Center (MIRCen), F-92265 Fontenay-aux-Roses, France

<sup>b</sup> CNRS CEA URA 2210, F-92265 Fontenay-aux-Roses, France

<sup>c</sup> CEA, iBiTecS, Service de Biologie Intégrative et Génétique Moléculaire (SBIGeM), F-91191 Gif-sur-Yvette  
Cedex, France

\* Corresponding author at: URA CEA-CNRS 2210, MIRCen, I2BM, CEA, 18 route du Panorama, BP 6, 92265 Fontenay-aux-Roses, France. Tel.: 0033-1-46549622; fax: 0033-1-46549116.

*E-mail address:* emmanuel.brouillet@cea.fr (E. Brouillet).

\*\* Alternate corresponding author at: iBiTecS, SBIGeM, Bâtiment 144, CEA Saclay, 91191 Gif-sur-Yvette Cedex, France. Tel.: 0033-1-69088865 ; fax: 0033-1-69084712.

*E-mail address:* michel.de.chaldee@cea.fr (M. de Chaldée).

<sup>1</sup> These authors were co-PIs for this work.

## **Abstract**

Genes selectively expressed in the striatum may be involved in the preferential vulnerability of striatal neurons to Huntington's disease (HD). We have previously shown that the expression of the Capucin (or Tmem90a) gene, which encodes a protein of unknown function, is enriched in the striatum and downregulated in HD models. Here, we investigated whether perturbations of Capucin expression could modify the neurotoxicity induced by the injection of a lentiviral vector encoding a short N-terminal fragment of mHtt into the mouse striatum. To this end, we engineered a Capucin knockout mouse line. Constitutive Capucin deficiency in knockout mice did not induce major anatomical abnormalities in the striatum and did not significantly modify the toxicity of the mHtt fragment. Consistent with this, lentiviral vector-mediated expression of Capucin in the striatum of adult Capucin knockout or wild-type mice had no marked effects on striatal vulnerability to the mHtt fragment. These results suggest that, despite its reduced striatal expression in HD, Capucin does not play a prominent role in mHtt toxicity.

*Keywords:* Huntington's disease; Capucin; Neurodegeneration; Basal ganglia; Striatal marker; Lentiviral vectors; Knockout mice

## 1. Introduction

Huntington's disease (HD) is a fatal neurodegenerative disorder with no effective treatment. It is caused by the expansion of a polyglutamine-encoding CAG repeat in the Huntingtin (Htt) gene (The Huntington's Disease Collaborative Research Group, 1993). Polyglutamine expansion confers toxicity on Htt, leading to a severe disturbance of major cellular functions, including transcription, protein degradation, mitochondrial energy production, calcium homeostasis and neurotransmission (Gil and Rego, 2008). The resulting mutant Huntingtin (mHtt) can aggregate, forming Ubiquitin-positive intranuclear inclusions in affected neurons (DiFiglia et al., 1997).

We have previously shown that normal brain aging is involved in the pathogenesis of HD (Diguët et al., 2009). However, the precise mechanisms of neurodegeneration remain to be determined. Although mHtt is ubiquitously expressed in the brain, the most prominent neuronal loss occurs in the striatum, a structure that plays a key role in movement initiation and cognitive functions (Vonsattel et al., 1985). This particular vulnerability may be conferred by factors that are enriched in the striatum, such as the GTP-binding protein Rhes, which has been shown to bind to mHtt and promote its cytotoxicity (Subramaniam et al., 2009), and D2 dopamine receptors, whose stimulation has been reported to increase neuronal sensitivity to mHtt by downregulating mitochondrial complex II expression (Charvin et al., 2005; Benchoua et al., 2008).

Other factors that are enriched in the striatum may also modulate the sensitivity of striatal neurons to mHtt and thereby provide novel perspectives for HD therapy. Recently, we compared the transcriptomes of 11 regions of the mouse brain and found ~ 100 genes that are preferentially expressed in the striatum (Brochier et al., 2008). Among them, the expression of the Capucin gene (also referred to as *Tmem90a*) is highly enriched in the striatum. Capucin is

a 237amino-acid protein of unknown function, with no recognizable motif other than two predicted carboxy-terminal transmembrane domains. It shares 45 % amino-acid identity with SynDIG1 (also known as Tmem90b), a protein involved in the development of AMPA receptor-containing synapses whose transcripts have been detected in the hippocampus and cerebellum (Kalashnikova et al., 2010). Notably, lower Capucin mRNA levels have been observed in the R6/2 transgenic mouse model of HD and in primary cultures of rat striatal neurons expressing a mutant fragment of human Htt than in the corresponding controls (de Chaldée et al., 2006). The same result has been obtained in the R6/1 transgenic mouse model (Desplats et al., 2006), suggesting a possible contribution of Capucin downregulation to the pathogenesis of HD.

In the present study, we examined whether perturbations in Capucin expression could modify striatal vulnerability to mHtt in mice. For this purpose, we used a model of mHtt toxicity developed in rats (de Almeida et al., 2002) and mice (Faideau et al., 2010) in which a lentiviral vector preferentially transducing neurons thanks to a vesicular stomatitis virus envelope and encoding a 171amino-acid N-terminal fragment of human Htt with either 82 (mutant) or 18 (wild-type) glutamine residues is injected into the striatum. To achieve a Capucin knockout, we engineered a novel mouse line that is described in the present paper. For intrastriatal Capucin re-expression in knockout mice and overexpression in wild-type mice, we co-injected a Capucin-encoding lentiviral vector with the Htt-fragment-expressing vector. Using lentiviral vectors encoding fluorescent proteins, we verified that co-injection, as co-transfection of cultured cells, resulted in co-expression in a majority of striatal neurons. For each approach, striatal vulnerability was assessed by histological evaluation: both the volume of the lesion induced by the mHtt fragment and the density of Ubiquitin-positive inclusions were quantified.

## 2. Methods

### 2.1. Animals

Adult male C57BL/6J mice (25 g each; Charles River, Saint-Germain-sur-l'Arbresle, France) and adult male Capucin knockout mice (generated as described below) were used for lentiviral infections. Animals were housed in a temperature-controlled room maintained on a 12 hr light/dark cycle. Food and water were available *ad libitum*. All experiments on mice were supervised by an investigator licensed by the Ministère de l'Agriculture, in conformity with French regulations and in strict accordance with the recommendations of the European Commission (86/609/EEC) concerning the care and use of laboratory animals.

### 2.2. Generation and genotyping of Capucin knockout mice

A 2,820-bp-long genomic fragment (nucleotides 86,018,915 to 86,021,734 on mouse chromosome 12) encompassing the entire Capucin coding region was replaced by a cassette containing the beta-galactosidase reporter gene (Fig. 1A) through homologous recombination in AT1 ES cells (Buchou et al., 2003), according to established procedures (Tessarollo, 2001). The 5' homology arm was 3,464-bp long (nucleotides 86,021,735 to 86,025,198 on mouse chromosome 12) and the 3' homology arm was 3,360-bp long (nucleotides 86,015,555 to 86,018,914 on mouse chromosome 12). Homozygous knockout mice and wild-type littermates were obtained on a mixed (25 % 129, 75 % C57BL/6N) background. Expression of the reporter gene was visualized by X-Gal staining according to standard protocols.



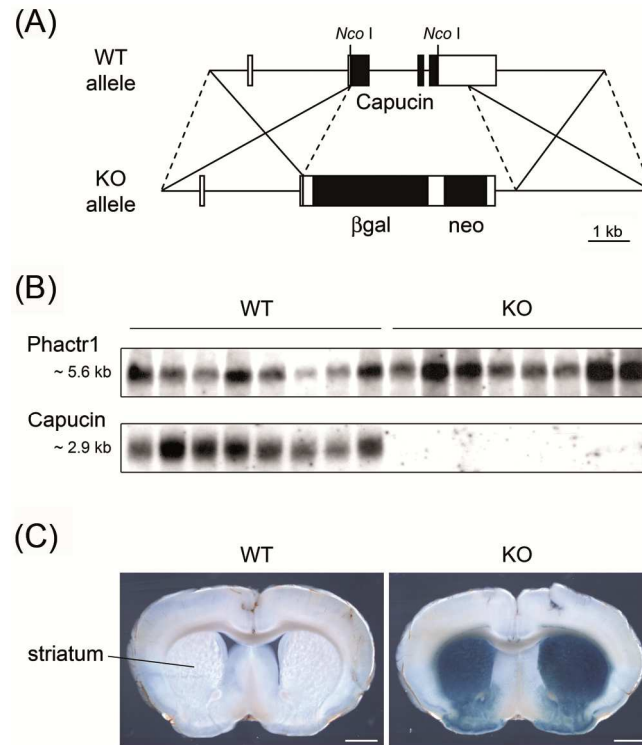


Figure 1

Fig. 1. Molecular characterization of *Capucin* knockout mice. (A) Wild-type (WT) and knockout (KO) *Capucin* alleles. Transcribed sequences are depicted by boxes and coding sequences are shown in black. Homologous strands are delimited by dotted lines and recombination events are represented by crossed lines.  $\beta$ gal, beta-galactosidase reporter gene; neo, neomycin selection gene. (B) Northern blot analysis of forebrain total RNA samples from eight wild-type (WT) mice and eight *Capucin* knockout (KO) littermates. The membrane was sequentially hybridized with two probes: a *Nco* I fragment of the *Capucin* transcribed DNA sequence (A) and a probe for the striatally enriched *Phactr1* transcript. Hybridization with the *Capucin* probe revealed no signal of any size with the samples from the knockout mice. (C) X-Gal staining of 300- $\mu$ m-thick coronal brain sections from adult wild-type (WT) mice and *Capucin* knockout (KO) littermates at the level of the striatum. Scale bar: 1 mm.

For genotyping, we developed a fast and convenient multiplex PCR assay using three primers (forward: 5' CCTGCCTGCGGAACCAT 3'; knockout reverse: 5' TGGGTAACGCCAGGGTTTT 3'; wild-type reverse: 5' GGCAGCAGTGGGTTCTGT 3') and two TaqMan probes (Life Technologies, Carlsbad, CA) labeled with distinct dyes (knockout probe: VIC-ATCCCGTCGTTTTACAACG-MGBNFQ; wild-type probe: 6FAM-AGAGCCTGAGTGAAC-MGBNFQ). Amplification occurs from the forward and knockout reverse primers for the knockout allele, and from the forward and wild-type reverse primers for the wild-type allele, yielding, respectively, a 68-bp-long product detected by the knockout probe and a 53-bp-long product detected by the wild-type probe. PCR was performed in a volume of 20 µl containing 1 x TaqMan Universal PCR Master Mix (Life Technologies, Carlsbad, CA), 500 nM forward primer, 400 nM knockout reverse primer, 400 nM wild-type reverse primer, 200 nM knockout probe, 200 nM wild-type probe and 30-40 ng of tail genomic DNA. Fluorescence was directly detected by an ABI prism 7300 or 7000 thermal cycler (Life Technologies, Carlsbad, CA) programmed for an initial incubation at 50°C for 2 min, followed by 95°C for 10 min, and 35 amplification cycles (95°C for 15 sec; 60°C for 1 min). Each individual was genotyped in duplicate.

### *2.3. Lentiviral vector construction, production and infection*

DNA sequences coding for green fluorescent protein (GFP), DsRed and C-terminal hemagglutinin (HA)-tagged mouse Capucin were cloned into the SIN-W-PGK lentiviral vector (de Almeida et al., 2002) to generate lenti-GFP, lenti-DsRed and lenti-Capucin-HA, respectively. The lentiviral vectors expressing a wild-type Htt fragment (lenti-Htt171-18Q), a mutant Htt fragment (lenti-Htt171-82Q) or beta-galactosidase (lenti-βgal) have been described previously (Diguët et al., 2009; Faideau et al., 2010). Viral particles were produced

as described elsewhere (Hottinger et al., 2000). The particle content of the viral batches was determined by ELISA for the p24 antigen (Gentaur, Paris, France). Lenti-Htt171-18Q and lenti-Htt171-82Q were used at a concentration of 150 ng/μl of p24, lenti-Capucin-HA and lenti-βgal at a concentration of 100 ng/μl of p24, lenti-GFP and lenti-DsRed at a concentration of 50 ng/μl of p24. A total volume of 2 μl of lentiviral suspension was injected into the mouse striatum as previously reported (Faideau et al., 2010), using the following stereotaxic coordinates: 1.0 mm anterior and 2.0 mm lateral to the bregma, at a depth of 2.7 mm from the dura, with the tooth bar set at 0.0 mm.

#### *2.4. Histological analyses*

After deep anesthesia by intraperitoneal injection of a sodium pentobarbital solution (50 μg per gram of body weight), mice were transcardially perfused with 100 ml of phosphate buffer containing 4 % paraformaldehyde at 8 ml/min. The brains were removed, post-fixed overnight in the same solution, then cryoprotected by immersion in a 15 % sucrose solution for 24 hours followed by immersion in a 30 % sucrose solution for another 24 hours. Free-floating 30-μm-thick serial coronal sections throughout the striatum were collected using a freezing sliding microtome (SM2400; Leica Microsystems, Wetzlar, Germany).

For immunohistochemistry, sections were treated with 0.3 % hydrogen peroxide for one hour, washed three times in phosphate-buffered saline (PBS), blocked in PBS containing 4.5 % normal goat serum for one hour, then incubated overnight at 4°C in PBS containing 3 % normal goat serum and one of the following antibodies: rabbit anti-DARPP-32 (Santa Cruz Biotechnology, Santa Cruz, CA; 1:1000), mouse anti-NeuN (Millipore, Molsheim, France; 1:200), rabbit anti-Ubiquitin (Wako Chemicals, Neuss, Germany; 1:1000), or mouse anti-HA (Covance, Princeton, NJ; 1:500). Sections were rinsed three times in PBS before

incubation with the appropriate anti-IgG biotinylated antibody (Vector Laboratories, Burlingame, CA) at a 1:5000 dilution for one hour. Staining was visualized by the addition of avidin-biotinylated peroxidase and incubation with DAB or VIP substrate (Vector Laboratories, Burlingame, CA) for one minute. For NeuN immunostaining, we used the M.O.M. immunodetection kit (Vector Laboratories, Burlingame, CA). Stained sections were mounted on microscopic slides.

The area of the striatal lesions resulting from lenti-Htt171-82Q infection was delineated manually by identifying the border of the lesion on each coronal brain section. The corresponding surface was calculated using the MCID image analysis software (InterFocus Imaging, Cambridge, UK). The volume of the striatal lesion was determined using the Cavalieri method (Diguët et al., 2009). The number of Ubiquitin-positive inclusions was quantified as previously described (Diguët et al., 2009) with the following modifications: the inter-section distance was 210  $\mu\text{m}$  (i.e. one in every seven sections was used) and observations were performed using a 10X objective on an Axioplan 2 Imaging microscope (Carl Zeiss, Le Pecq, France) equipped with a motorized stage and an automated image acquisition and analysis system (Mercator software, Explora Nova, La Rochelle, France). With this set-up, objects with an apparent cross-sectional area of over 5  $\mu\text{m}^2$  could be reliably detected.

For GFP fluorescence analysis, equidistant 30- $\mu\text{m}$ -thick coronal brain sections (intersection distance: 210  $\mu\text{m}$ ) were randomly sampled throughout the striatum. Observations were performed using a 10X objective on a motorized Axioplan 2 Imaging microscope (Carl Zeiss, Le Pecq, France) equipped with an X-Cite 120PC Q fluorescence excitation system (Lumen Dynamics, Mississauga, Canada) allowing controlled excitation intensity and homogeneity throughout the field of view. Images were automatically acquired and analyzed using Fluo'Up and Mercator softwares (Explora Nova, La Rochelle, France). Care was taken

to optimize image acquisition and avoid image saturation. On each section, the fluorescence level of each lenti-GFP-infected striatum was determined by subtracting the background fluorescence level measured outside the infection site from the mean fluorescence level within the infected area. A mean fluorescence level was then calculated across sections for each striatum.

## *2.5. Real-time quantitative RT-PCR and Northern blot analysis*

Adult mice were deeply anesthetized by intraperitoneal injection of a sodium pentobarbital solution (50 µg per gram of body weight) before decapitation. The brains were immediately removed and positioned in a coronal brain matrix (Ted Pella, Redding, CA). For the quantification of SynDIG1 mRNA levels in knockout mice, the striatum was dissected out using a scalpel from 1-mm-thick brain sections. Total RNA extraction and real-time quantitative RT-PCR (qRT-PCR) were carried out according to standard procedures, using the primers 5' GAGGAGGAGAGCGATGGG 3' and 5' CAGGGTGAGGAAGGCGTC 3' for Capucin, and 5' GTTGCCCTCATTGCCTACCT 3' and 5' TGTCTCTTCTCTGCCCTCGTACA 3' for SynDIG1. For Northern blot analysis, a 5-mm-thick block containing the entire striatum (between grooves 3 and 8 of the matrix, groove 1 being the most rostral; Brochier et al., 2008) was removed from each brain and processed for mRNA extraction. Twenty-five-microgram total RNA samples were denatured in a glyoxal-dimethylsulfoxide solution, electrophoresed on a 1 % agarose gel, and then transferred onto a Hybond-XL nylon membrane (GE Healthcare, Little Chalfont, UK). Hybridization was carried out at 65°C in Rapid-Hyb buffer (GE Healthcare, Little Chalfont, UK) using an [ $\alpha$ -<sup>32</sup>P]dCTP random prime-labeled (Prime-It II; Agilent Technologies, Santa Clara, CA) 706-bp-long *Nco* I fragment of the Capucin transcribed DNA sequence as a probe (Fig. 1A).

An [ $\alpha$ - $^{32}$ P]dCTP random prime-labeled 554-bp-long *Nco* I fragment of the striatally enriched Phactr1 transcribed DNA sequence (de Chaldée et al., 2006; Allen et al., 2004) was used to control for RNA input.

### 3. Results

#### 3.1. Development of a Capucin knockout mouse line

To create a new resource for studying Capucin function, we developed a Capucin knockout mouse line. Homozygous knockout mice were obtained at a similar ratio to homozygous wild-type littermates (of the 745 mice genotyped so far 167 were homozygous knockout mice, 188 were homozygous wild-type mice and 390 were heterozygous mice). Knockout mice are apparently healthy and their behavior is normal at the gross level. Both males and females are fertile. The brain shows no overt anatomical abnormalities. In particular, morphometric analysis indicates a striatal volume similar to that of wild-type littermates ( $6.152 \pm 0.310 \text{ mm}^3$  in knockout mice,  $6.413 \pm 0.334 \text{ mm}^3$  in wild-type littermates;  $n = 10$  knockout and 10 wild-type adult mice). Knockout mice lack the entire coding region of the Capucin gene (Fig. 1A). The absence of transcripts containing even a part of the Capucin coding region in the striatum of knockout mice was confirmed by Northern blot analysis using the coding sequence as a probe (Fig. 1B). To assess whether the lack of a conspicuous phenotype could be attributable to sequence redundancy, we measured striatal mRNA levels of SynDIG1, the closest paralog of Capucin, using qRT-PCR. No significant difference was observed between genotypes (SynDIG1 mRNA levels in Capucin-knockout striata were  $88.1 \pm 10.7 \%$  of those in wild-type samples;  $p = 0.209$ , unpaired Student's *t*-test,  $n = 7$

knockout and 7 wild-type mice), arguing against functional compensation of Capucin deficiency by SynDIG1 overexpression.

We took advantage of the beta-galactosidase transgene to investigate the spatiotemporal pattern of expression driven by the Capucin gene promoter. As expected, X-Gal staining was intense in the striatum of adult knockout animals (Fig. 1C). During development, X-Gal staining was first observed in the medulla oblongata of 15.5 days post-coitum (dpc) embryos, and in the olfactory bulbs and discrete regions of the diencephalon, mesencephalon and rhombencephalon of 16.5 dpc embryos, predicting Capucin gene expression in nuclei of the adult thalamus, midbrain and hindbrain, in particular in the trigeminal, facial and hypoglossal nuclei (de Chaldée et al., 2006; Lein et al., 2007). Staining in the striatum was first clearly detected five days after birth. We also examined adult peripheral tissues and noticed stained patches in the pancreas. However, we observed neither overt polyuria nor hyperglycemia in Capucin knockout mice (serum concentrations of glucose were  $1.38 \pm 0.09$  g/l in adult knockout mice [ $n = 22$ ] and  $1.41 \pm 0.10$  g/l in wild-type littermates [ $n = 24$ ]).

### *3.2. Effect of Capucin knockout on striatal vulnerability to a mutant Htt fragment*

Adult Capucin knockout mice and wild-type littermates were intrastrially injected with a mixed suspension of lentiviral vectors encoding a wild-type or mutant Htt fragment (lenti-Htt171-18Q or lenti-Htt171-82Q, respectively) and GFP (lenti-GFP). GFP fluorescence was used to monitor neuronal survival, as loss of fluorescence has been shown to be correlated with neuronal death (Arrasate and Finkbeiner, 2005). Histological evaluation was performed six weeks post-infection. At this time point, GFP fluorescence was much fainter in lenti-Htt171-82Q- than in lenti-Htt171-18Q-infected striata (Fig. 2A), consistent with the

expected toxicity of the mHtt fragment. Immunostaining for DARPP-32, a well-known marker of striatal neurons, and for NeuN, a widely used neuronal marker, revealed the presence of a lesion around the injection site in lenti-Htt171-82Q- but not in lenti-Htt171-18Q-infected striata (Fig. 2B and C; Fig. 4B and C). Brain sections were also examined for Ubiquitin-positive inclusions. They were detected in lenti-Htt171-82Q- but not in lenti-Htt171-18Q-infected striata (Figs. 2D and 4D).

Comparison between knockout and wild-type lenti-Htt171-82Q-infected striata indicated no significant differences in the loss of fluorescence (Fig. 2A), the lesion volume as determined by DARPP-32 (Fig. 2B) and NeuN (Fig. 2C) immunostaining, or the density of Ubiquitin-positive inclusions (Fig. 2D).



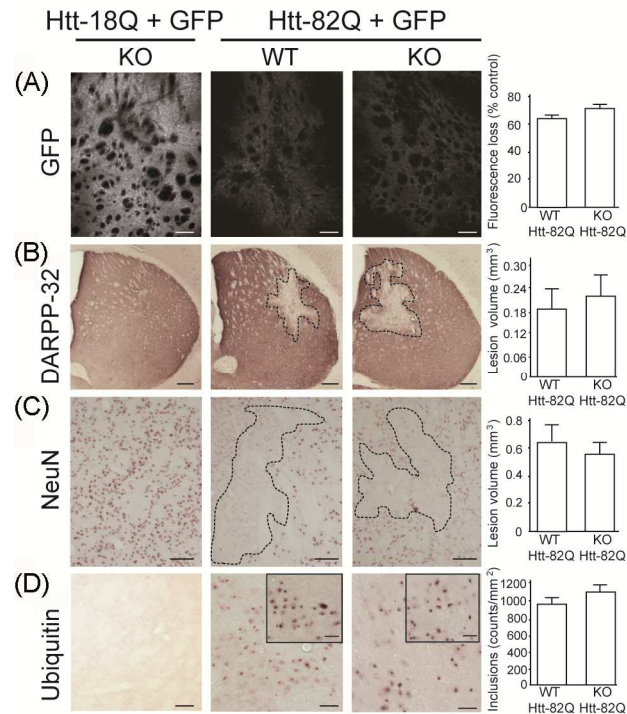


Figure 2

Fig. 2. Effect of Capucin knockout on striatal vulnerability to a mutant Htt fragment. Adult male Capucin knockout (KO) mice ( $n = 10$ ) and wild-type (WT) littermates ( $n = 10$ ) received a bilateral intrastratial injection of lenti-GFP. Lenti-Htt171-18Q was co-injected into the left striatum and lenti-Htt171-82Q into the right striatum of each mouse. Six weeks after infection, brains were processed for GFP fluorescence analysis (A) and DARPP-32 (B), NeuN (C) and Ubiquitin (D) immunostaining. The loss of GFP fluorescence (A) was expressed as the difference in fluorescence levels between lenti-Htt171-18Q- and lenti-Htt171-82Q-infected striata divided by the fluorescence level of the lenti-Htt171-18Q-infected striatum. The loss of GFP fluorescence (A), the volume of the lesion (dotted line), as determined by DARPP-32 (B) and NeuN (C) immunostaining, and the density of Ubiquitin-positive inclusions (D) are displayed as means  $\pm$  SEM. Differences between lenti-Htt171-82Q-infected knockout and wild-type striata were assessed by an unpaired Student's *t*-test. Scale bars: 0.2 mm (A), 0.5 mm (B), 0.1 mm (C), 0.05 mm (D), 0.01 mm (D, insets).

### *3.3. Effect of Capucin re-expression in knockout striata on vulnerability to a mutant Htt fragment*

Testing the effects of Capucin re-expression in knockout striata on vulnerability to mHtt requires co-expression of recombinant Capucin and the mutant Htt fragment in the same striatal neurons. To verify that the co-injection of two lentiviral vectors leads to actual co-expression in a majority of cells, we delivered a mixed suspension of lenti-GFP and a lentiviral vector encoding DsRed (lenti-DsRed) into wild-type mouse striata. Six weeks after infection, we observed colocalization of both fluorescent proteins in most cells, indicating that co-injection produced effective co-expression of both transgenes in the same cells.

We therefore injected a mixed suspension of lenti-Htt171-82Q and a lentiviral vector encoding HA-tagged Capucin (lenti-Capucin-HA) into the striata of knockout mice. Histological evaluation was carried out six weeks post-infection. Recombinant Capucin expression was verified in all lenti-Capucin-HA-infected striata using anti-HA immunohistochemistry (Fig. 3A). DARPP-32 (Fig. 3B) and NeuN (Fig. 3C) immunostaining revealed no significant difference in lesion volumes between lenti-Capucin-HA- and lenti-GFP-infected knockout striata. No difference in the density of Ubiquitin-positive inclusions could be detected either (Fig. 3D).

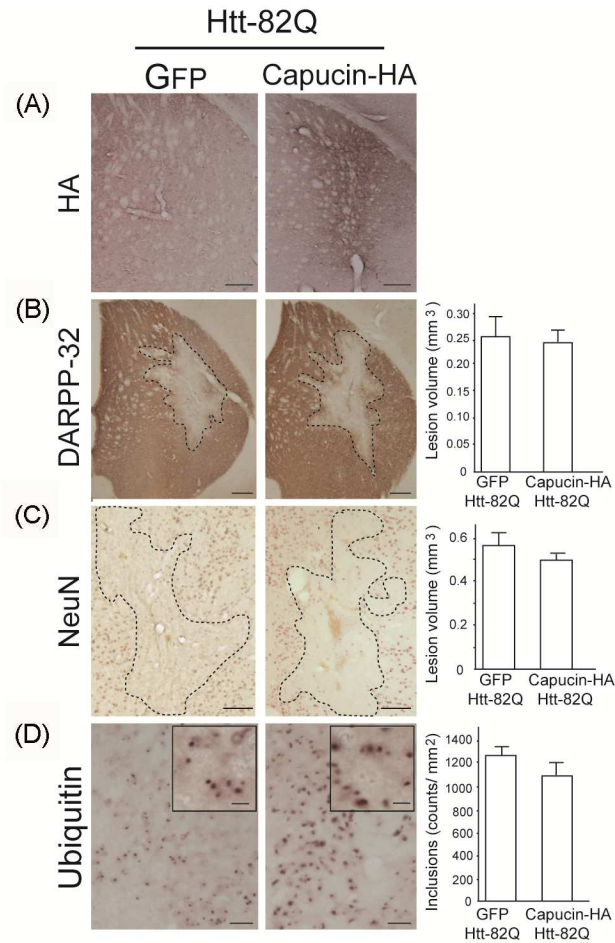


Figure 3

Fig. 3. Effect of Capucin re-expression in knockout striata on vulnerability to a mutant Htt fragment. Adult male Capucin knockout mice ( $n = 14$ ) received a bilateral intrastriatal injection of lenti-Htt171-82Q. Lenti-GFP was co-injected into the left striatum and lenti-Capucin-HA into the right striatum of each mouse. Six weeks after infection, brains were processed for HA (A), DARPP-32 (B), NeuN (C) and Ubiquitin (D) immunostaining. The volume of the lesion (dotted line), as determined by DARPP-32 (B) and NeuN (C) immunostaining, and the density of Ubiquitin-positive inclusions (D) are displayed as means  $\pm$  SEM. Differences between striata co-infected with lenti-Capucin-HA and lenti-Htt171-82Q and striata co-infected with lenti-GFP and lenti-Htt171-82Q were assessed by a paired Student's *t*-test. Scale bars: 0.2 mm (A), 0.5 mm (B), 0.1 mm (C), 0.05 mm (D), 0.01 mm (D, insets).

### *3.4. Effect of Capucin overexpression in wild-type striata on vulnerability to a mutant Htt fragment*

We next examined whether increasing Capucin expression levels in adult wild-type mice could modify the toxicity of a mutant Htt fragment. Overexpression of Capucin in mouse striata was achieved by stereotaxic injection of lenti-Capucin-HA. Overexpression efficiency was measured six weeks post-infection by qRT-PCR. The results show that Capucin mRNA levels in extracts from lenti-Capucin-HA-infected striata ( $n = 6$ ) were  $200 \pm 26.6 \%$  of those in extracts from striata infected with a beta-galactosidase-encoding lentiviral vector (lenti- $\beta$ gal;  $n = 6$ ). This difference is statistically significant ( $p = 0.011$ , paired Student's  $t$ -test).

A mixed suspension of lenti-Htt171-82Q and lenti-Capucin-HA was then delivered into the striatum of wild-type mice. Expression of recombinant Capucin in the striatum of infected mice was checked six weeks post-infection by anti-HA immunohistochemistry (Fig. 4A). At this time point, neither DARPP-32 (Fig. 4B) nor NeuN (Fig. 4C) immunostaining revealed any significant change in lesion volume in lenti-Capucin-HA-infected striata as compared to striata infected with lenti- $\beta$ gal. Notably though, the density of Ubiquitin-positive inclusions was somewhat lower in lenti-Capucin-HA-infected striata (Fig. 4D).

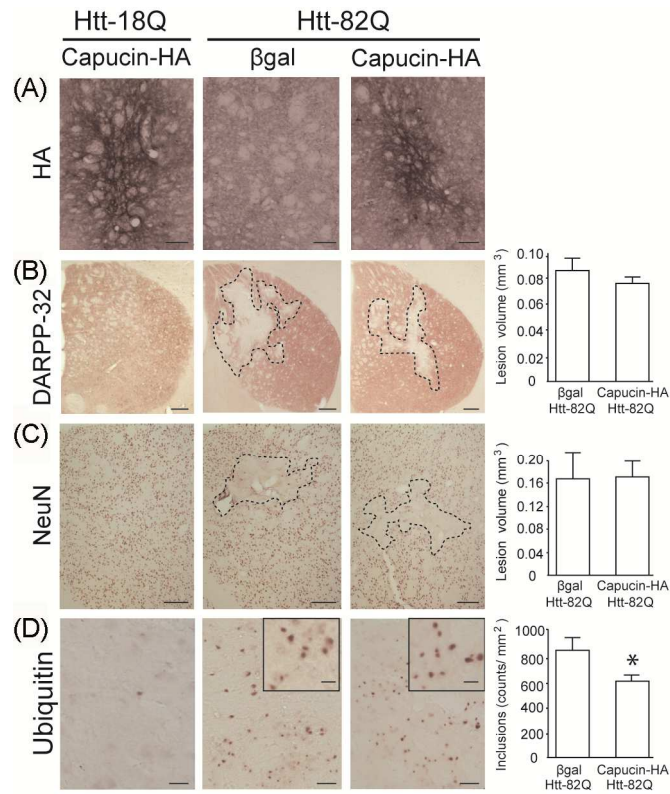


Figure 4

Fig. 4. Effect of Capucin overexpression in wild-type striata on vulnerability to a mutant Htt fragment. Adult male C57BL/6J mice received a bilateral intrastratial injection of either lenti-Capucin-HA ( $n = 10$ ) or lenti- $\beta$ gal ( $n = 10$ ). Lenti-Htt171-18Q was co-injected into the left striatum and lenti-Htt171-82Q into the right striatum of each mouse. Six weeks after infection, brains were processed for HA (A), DARPP-32 (B), NeuN (C) and Ubiquitin (D) immunostaining. The volume of the lesion (dotted line), as determined by DARPP-32 (B) and NeuN (C) immunostaining, and the density of Ubiquitin-positive inclusions (D) are displayed as means  $\pm$  SEM. Differences between striata co-infected with lenti-Capucin-HA and lenti-Htt171-82Q and striata co-infected with lenti- $\beta$ gal and lenti-Htt171-82Q were assessed by an unpaired Student's  $t$ -test (\*  $p = 0.042$ ). Scale bars: 0.2 mm (A), 0.5 mm (B), 0.1 mm (C), 0.05 mm (D), 0.01 mm (D, insets).

## 4. Discussion

The mechanisms underlying the preferential degeneration of the striatum in HD remain unknown. Using a combination of *in vivo* approaches, we tested the hypothesis that the Capucin gene, whose expression is highly enriched in the striatum and significantly downregulated in transgenic HD models (de Chaldée et al., 2006), may play a role in the striatal sensitivity to mHtt.

Our results show that neither knocking out nor increasing Capucin expression leads to a major modification of the toxicity of an N-terminal fragment of mHtt in mice. The effect of the continuous expression of mHtt in our model has been characterized by the loss of expression of two neuronal markers, DARPP-32 and NeuN. The loss of NeuN immunoreactivity is associated with actual neuronal degeneration and dysfunction, as assessed using markers of cell death (de Almeida and al., 2002) and stereological cell counts (Diguet et al., 2009). Thus, the present results indicate that modifying Capucin expression has no effect on the cell death induced by the mHtt fragment.

A Capucin knockout mouse line was generated for the present study. The absence of the entire Capucin coding sequence in knockout mice was ascertained at both the genomic and transcript levels. These mice show no obvious abnormalities. Although we verified that Capucin knockout was not compensated for by the overexpression of its closest paralog, SynDIG1, functional compensation by proteins with unrelated sequences cannot be ruled out. Being fertile and apparently healthy, Capucin knockout mice can be conveniently used in a variety of behavioral tests, including in late adulthood, and can be crossed with mice from other lines.

The effects of Capucin overexpression in wild-type striata and re-expression in knockout striata on mHtt-fragment-induced toxicity were assessed by intrastriatal co-injection

of lentiviral vectors expressing Capucin and a mutant Htt fragment. Actual co-infection of striatal neurons with both vectors is supported by experiments in which lentiviral vectors encoding GFP and DsRed were co-injected into wild-type mouse striata, resulting in colocalization of both fluorescent proteins in the same cells. Using similar co-infection approaches, we have identified a number of modifiers of the toxicity induced by the mHtt fragment in rats or mice, including transcription factor CA150 (Arango et al., 2006), chaperone proteins (Perrin et al., 2007), and signaling proteins of the JNK pathway (Perrin et al., 2009).

At the transcriptional level, Capucin expression is strongly downregulated in HD models (de Chaldée et al., 2006; Desplats et al., 2006). The downregulation of a number of transcripts has been reported in the striatum of HD models (Brochier et al., 2008; Desplats et al., 2006) and patients (Hodges et al., 2006), and may result from the functional disruption of striatum-specific transcriptional activators such as Bcl11b (Desplats et al., 2008). Depending on the transcript, this phenomenon could be protective, detrimental or have no effect on striatal neurons. For instance, the downregulation of the signalling protein CalDag-GEFI has been shown to be neuroprotective in HD models (Crittenden et al., 2010). On the contrary, the downregulation of the sodium channel subunit beta-4 (Scn4b), which can promote neurite outgrowth in cultured cells (Oyama et al., 2006), is expected to increase neuronal vulnerability in HD conditions. Likewise, overexpression of the HD-downregulated mitogen- and stress-activated protein kinase 1 (MSK-1) in model rats has been reported to be neuroprotective, while MSK-1 knockout mice show spontaneous striatal atrophy as they age and are more susceptible to 3-nitropropionic acid-induced striatal lesions (Martin et al., 2011). On the other hand, overexpression of the neuronal calcium sensor Hippocalcin in primary cultures of striatal neurons has no effect on neurotoxicity in cellular models of HD, despite its markedly reduced expression in patients (Rudinskiy et al., 2009). Similarly, the

downregulation of Capucin expression in HD conditions may have no effect on striatal vulnerability, further emphasizing that reduced striatal expression of a gene in HD does not *a fortiori* imply that the gene product plays a role in the disease process.

In the present work, we provide the first *in vivo* investigation of the possible involvement of Capucin in the striatal sensitivity to mHtt. Our results suggest that Capucin expression levels have no impact on mHtt toxicity.

## Disclosure statement

The authors have no conflict of interest to declare.

## Acknowledgements

The authors thank Nicolas Caudy (CEA, iBiTecS, SPI), Hélène Humbertclaude (CEA, iBiTecS, SBIGeM) and Sylvie Jounier (CEA, iBiTecS, SBIGeM) for technical assistance. They are grateful to Muriel Vernet (CEA, iRCM, SCSR) for providing AT1 ES cells and to Catherine Bourcier (CEA, DSM Saclay, UPSE, LABM) for supervising glycemia analysis. We also thank Philippe Hantraye (CEA, I2BM, MIRCen, URA CEA-CNRS 2210) for his constant support and encouragement.

The research leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. HEALTH-F5-2008-222925. LG was supported by the Neuropôle de Recherche Francilien and the Fondation pour la Recherche Médicale. NL was the recipient of a Ph.D. fellowship from the Ministère de l'Education Nationale, de la Recherche et de la Technologie. We also received financial support from the IMAGEN European integrated project (EUR617037286).



## References

- Allen, P.B., Greenfield, A.T., Svenningsson, P., Haspeslagh, D.C., Greengard, P., 2000. Phactrs 1-4: A family of protein phosphatase 1 and actin regulatory proteins. *Proc. Natl. Acad. Sci. U. S. A.* 101, 7187-7192.
- Arango, M., Holbert, S., Zala, D., Brouillet, E., Pearson, J., Régulier, E., Thakur, A.K., Aebischer, P., Wetzel, R., Déglon, N., Néri, C., 2006. CA150 expression delays striatal cell death in overexpression and knock-in conditions for mutant huntingtin neurotoxicity. *J. Neurosci.* 26, 4649-4659.
- Arrasate, M., Finkbeiner, S., 2005. Automated microscope system for determining factors that predict neuronal fate. *Proc. Natl. Acad. Sci. U. S. A.* 102, 3840-3845.
- Benchoua, A., Trioulier, Y., Diguët, E., Malgorn, C., Gaillard, M.C., Dufour, N., Elalouf, J.M., Krajewski, S., Hantraye, P., Déglon, N., Brouillet, E., 2008. Dopamine determines the vulnerability of striatal neurons to the N-terminal fragment of mutant huntingtin through the regulation of mitochondrial complex II. *Hum. Mol. Genet.* 17, 1446-1456.
- Brochier, C., Gaillard, M.C., Diguët, E., Caudy, N., Dossat, C., Ségurens, B., Wincker, P., Roze, E., Caboche, J., Hantraye, P., Brouillet, E., Elalouf, J.M., de Chaldée M., 2008. Quantitative gene expression profiling of mouse brain regions reveals differential transcripts conserved in human and affected in disease models. *Physiol. Genomics* 33, 170-179.
- Buchou, T., Vernet, M., Blond, O., Jensen, H.H., Pointu, H., Olsen, B.B., Cochet, C., Issinger, O.G., Boldyreff, B., 2003. Disruption of the regulatory beta subunit of protein kinase CK2 in mice leads to a cell-autonomous defect and early embryonic lethality. *Mol. Cell. Biol.* 23, 908-915.

- Charvin, D., Vanhoutte, P., Pagès, C., Borrelli, E., Caboche, J., 2005. Unraveling a role for dopamine in Huntington's disease: the dual role of reactive oxygen species and D2 receptor stimulation. *Proc. Natl. Acad. Sci. U. S. A.* 102, 12218-12223.
- Crittenden, J.R., Dunn, D.E., Merali, F.I., Woodman, B., Yim, M., Borkowska, A.E., Frosch, M.P., Bates, G.P., Housman, D.E., Lo, D.C., Graybiel, A.M., 2010. CalDAG-GEFI down-regulation in the striatum as a neuroprotective change in Huntington's disease. *Hum. Mol. Genet.* 19, 1756-1765.
- de Almeida, L.P., Ross, C.A., Zala, D., Aebischer, P., Déglon, N., 2002. Lentiviral-mediated delivery of mutant huntingtin in the striatum of rats induces a selective neuropathology modulated by polyglutamine repeat size, huntingtin expression levels, and protein length. *J. Neurosci.* 22, 3473-3483.
- de Chaldée, M., Brochier, C., Van de Vel, A., Caudy, N., Luthi-Carter, R., Gaillard, M.C., Elalouf, J.M., 2006. Capucin: a novel striatal marker down-regulated in rodent models of Huntington disease. *Genomics* 87, 200-207.
- Desplats, P.A., Kass, K.E., Gilmartin, T., Stanwood, G.D., Woodward, E.L., Head, S.R., Sutcliffe, J.G., Thomas, E.A., 2006. Selective deficits in the expression of striatal-enriched mRNAs in Huntington's disease. *J. Neurochem.* 96, 743-757.
- Desplats, P.A., Lambert, J.R., Thomas, E.A., 2008. Functional roles for the striatal-enriched transcription factor, Bcl11b, in the control of striatal gene expression and transcriptional dysregulation in Huntington's disease. *Neurobiol. Dis.* 31, 298-308.
- DiFiglia, M., Sapp, E., Chase, K.O., Davies, S.W., Bates, G.P., Vonsattel, J.P., Aronin, N., 1997. Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science* 277, 1990-1993.

- Diguet, E., Petit, F., Escartin, C., Cambon, K., Bizat, N., Dufour, N., Hantraye, P., Déglon, N., Brouillet, E., 2009. Normal aging modulates the neurotoxicity of mutant huntingtin. *PLoS One* 4, e4637.
- Faideau, M., Kim, J., Cormier, K., Gilmore, R., Welch, M., Auregan, G., Dufour, N., Guillermier, M., Brouillet, E., Hantraye, P., Déglon, N., Ferrante, R.J., Bonvento, G., 2010. In vivo expression of polyglutamine-expanded huntingtin by mouse striatal astrocytes impairs glutamate transport: a correlation with Huntington's disease subjects. *Hum. Mol. Genet.* 19, 3053-3067.
- Gil, J.M., Rego, A.C., 2008. Mechanisms of neurodegeneration in Huntington's disease. *Eur. J. Neurosci.* 27, 2803-2820.
- Hodges, A., Strand, A.D., Aragaki, A.K., Kuhn, A., Sengstag, T., Hughes, G., Elliston, L.A., Hartog, C., Goldstein, D.R., Thu, D., Hollingsworth, Z.R., Collin, F., Synek, B., Holmans, P.A., Young, A.B., Wexler, N.S., Delorenzi, M., Kooperberg, C., Augood, S.J., Faull, R.L., Olson, J.M., Jones, L., Luthi-Carter, R., 2006. Regional and cellular gene expression changes in human Huntington's disease brain. *Hum. Mol. Genet.* 15, 965-977.
- Hottinger, A.F., Azzouz, M., Déglon, N., Aebischer, P., Zurn, A.D., 2000. Complete and long-term rescue of lesioned adult motoneurons by lentiviral-mediated expression of glial cell line-derived neurotrophic factor in the facial nucleus. *J. Neurosci.* 20, 5587-5593.
- Kalashnikova, E., Lorca, R.A., Kaur, I., Barisone, G.A., Li, B., Ishimaru, T., Trimmer, J.S., Mohapatra, D.P., Díaz, E., 2010. SynDIG1: an activity-regulated, AMPA- receptor-interacting transmembrane protein that regulates excitatory synapse development. *Neuron* 65, 80-93.

Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A.F., Boguski, M.S., Brockway, K.S., Byrnes, E.J., Chen, L., Chen, L., Chen, T.M., Chin, M.C., Chong, J., Crook, B.E., Czaplinska, A., Dang, C.N., Datta, S., Dee, N.R., Desaki, A.L., Desta, T., Diep, E., Dolbeare, T.A., Donelan, M.J., Dong, H.W., Dougherty, J.G., Duncan, B.J., Ebbert, A.J., Eichele, G., Estin, L.K., Faber, C., Facer, B.A., Fields, R., Fischer, S.R., Fliss, T.P., Frensley, C., Gates, S.N., Glattfelder, K.J., Halverson, K.R., Hart, M.R., Hohmann, J.G., Howell, M.P., Jeung, D.P., Johnson, R.A., Karr, P.T., Kawal, R., Kidney, J.M., Knapik, R.H., Kuan, C.L., Lake, J.H., Laramee, A.R., Larsen, K.D., Lau, C., Lemon, T.A., Liang, A.J., Liu, Y., Luong, L.T., Michaels, J., Morgan, J.J., Morgan, R.J., Mortrud, M.T., Mosqueda, N.F., Ng, L.L., Ng, R., Orta, G.J., Overly, C.C., Pak, T.H., Parry, S.E., Pathak, S.D., Pearson, O.C., Puchalski, R.B., Riley, Z.L., Rockett, H.R., Rowland, S.A., Royall, J.J., Ruiz, M.J., Sarno, N.R., Schaffnit, K., Shapovalova, N.V., Sivasay, T., Slaughterbeck, C.R., Smith, S.C., Smith, K.A., Smith, B.I., Sodt, A.J., Stewart, N.N., Stumpf, K.R., Sunkin, S.M., Sutram, M., Tam, A., Teemer, C.D., Thaller, C., Thompson, C.L., Varnam, L.R., Visel, A., Whitlock, R.M., Wohnoutka, P.E., Wolkey, C.K., Wong, V.Y., Wood, M., Yaylaoglu, M.B., Young, R.C., Youngstrom, B.L., Yuan, X.F., Zhang, B., Zwingman, T.A., Jones, A.R., 2007. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445, 168-176.

Martin, E., Betuing, S., Pagès, C., Cambon, K., Auregan, G., Deglon, N., Roze, E., Caboche, J., 2011. Mitogen- and stress-activated protein kinase 1-induced neuroprotection in Huntington's disease: role on chromatin remodeling at the PGC-1-alpha promoter. *Hum. Mol. Genet.* 20, 2422-2434.

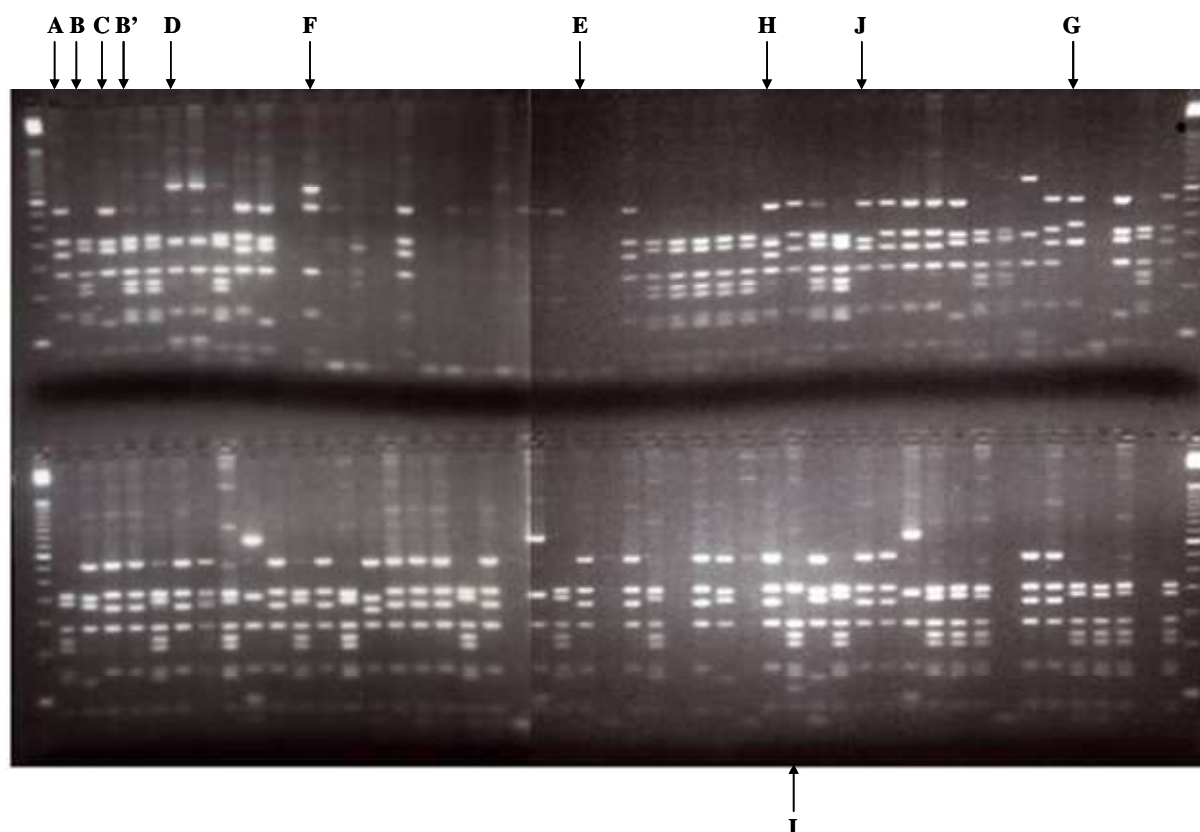
Oyama, F., Miyazaki, H., Sakamoto, N., Becquet, C., Machida, Y., Kaneko, K., Uchikawa, C., Suzuki, T., Kurosawa, M., Ikeda, T., Tamaoka, A., Sakurai, T., Nukina, N., 2006.

- Sodium channel beta4 subunit: down-regulation and possible involvement in neuritic degeneration in Huntington's disease transgenic mice. *J. Neurochem.* 98, 518-529.
- Perrin, V., Dufour, N., Raoul, C., Hassig, R., Brouillet, E., Aebischer, P., Luthi-Carter, R., Déglon, N., 2009. Implication of the JNK pathway in a rat model of Huntington's disease. *Exp. Neurol.* 215, 191-200.
- Perrin, V., Régulier, E., Abbas-Terki, T., Hassig, R., Brouillet, E., Aebischer, P., Luthi-Carter, R., Déglon, N., 2007. Neuroprotection by Hsp104 and Hsp27 in lentiviral-based rat models of Huntington's disease. *Mol. Ther.* 15, 903-911.
- Rudinskiy, N., Kaneko, Y.A., Beesen, A.A., Gokce, O., Régulier, E., Déglon, N., Luthi-Carter, R., 2009. Diminished hippocalcin expression in Huntington's disease brain does not account for increased striatal neuron vulnerability as assessed in primary neurons. *J. Neurochem.* 111, 460-472.
- Subramaniam, S., Sixt, K.M., Barrow, R., Snyder, S.H., 2009. Rhes, a striatal specific protein, mediates mutant-huntingtin cytotoxicity. *Science* 324, 1327-1330.
- Tessarollo, L., 2001. Manipulating mouse embryonic stem cells. *Methods Mol. Biol.* 158, 47-63.
- The Huntington's Disease Collaborative Research Group, 1993. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 72, 971-983.
- Vonsattel, J.P., Myers, R.H., Stevens, T.J., Ferrante, R.J., Bird, E.D., Richardson, E.P. Jr., 1985. Neuropathological classification of Huntington's disease. *J. Neuropathol. Exp. Neurol.* 44, 559-577.

## 4. Production d'anticorps contre la Capucine

Afin de poursuivre la caractérisation de la Capucine une production d'anticorps a été initiée. Pour ce faire différentes stratégies ont été mises en place (production d'anticorps anti-peptides de poule ; production d'anticorps contre un fragment recombinant de Capucine chez le lapin), mais sans succès. Nous avons donc décidé de nous orienter vers la méthode de production d'anticorps monoclonaux *in vitro* par « *phage display* ». Cette méthode est basée sur la capacité particulière des bactériophages d'exposer à leur surface des protéines qu'ils expriment. On se sert de cette propriété pour exposer des fragments variables (scFv - *Single-chain variable fragment*) d'immunoglobulines. Les scFv sont des protéines de fusion entre les régions variables des chaînes lourde ( $V_H$ ) et légère ( $V_L$ ) d'immunoglobulines isolées à partir de lymphocytes périphériques naïfs et combinées de façon aléatoire par PCR. La protéine d'enveloppe du bactériophage pIII, qui va être fusionnée sur l'extrémité N-terminale des scFv, permet leur exposition à la surface des phages (Marks et al., 1992). Un tag Myc est également fusionné aux scFv. L'avantage majeur de cette technique est qu'elle permet non seulement de contourner l'étape d'immunisation d'un animal, mais aussi de conserver les clones obtenus grâce à la production d'une banque de phages.

Une banque de phages exprimant différents scFv (Marks et al., 1991) a été utilisée pour effectuer cette expérience. Des clones candidats, qui expriment des scFv liant un fragment N-terminal de Capucine de 150 résidus produit dans *Escherichia coli*, ont été obtenus après trois tours de sélection (d'après le protocole décrit par Schier et al., 1995). Nous avons analysé la diversité des scFv sélectionnés par *fingerprint* (amplification par PCR des gènes codants les scFv et digestion des amplicons par BstNI comme décrit par Marks et al., 1991). Onze profils de digestion différents ont été observés sur 96 scFv (Figure 21). Nous avons caractérisé la séquence de certains d'entre eux (Figure 22).



**Figure 21 : Fingerprint des gènes de scFv.**

Les gènes codant les scFv ont été amplifiés par PCR et les amplicons ont été digérés par l'enzyme de restriction BstNI. Les fragments d'ADN digéré ont été séparés par électrophorèse sur un gel d'agarose à 3,5%. Le marqueur de taille utilisé est le 50bp ladder. Onze profils de digestion différents sont nommés de A à J. Pour des raisons de mauvaise qualité d'image, le profil E a été redéposé une nouvelle fois (résultat non montré).

| Heavy chains |                            |          |                   |                  |                  |                         |               |       |        |
|--------------|----------------------------|----------|-------------------|------------------|------------------|-------------------------|---------------|-------|--------|
| Profil       | IMGT FR1                   | CDR1     | FR2               | CDR2             | FR3              | CDR3                    | FR4           |       |        |
| A            | QVQLVESGGGVVQPGRSLRLSCAAS  | GFRFRSYG | MHWVRQAPGKGLEWVAN | IKPDGSEK         | YYGDSVKGRFTISRDN | AKNSLYLQLSSLRAEDTAVYYC  | ARNRGYATLDY   | WGQGT | LVTVSS |
| B            | QVQLVESGGGLVQPGGSLRLSCAAS  | GFTFSNHG | MNWVRQAPGKGLEWLSY | IGGSGGSI         | RNADSVKGRFTISRDN | AKNSLSLQMNSLRAEDTAMYYC  | ARGGGRLLDY    | WGQGT | LVTVSS |
| B'           | QVQLVESGGGLVQPGGSLRLSCAAS  | GFTFSNHG | MNWVRQAPGKGLEWLS  | LVVLPVPP         | ...              |                         |               |       |        |
| C            | QVQLVESGGGVVQPGRSLRLSCADS  | GFTFSNDW | MTWVRQAPGKGLEWVAI | INQDGNEK         | YYLDSVRGRFTISRDN | ARNLSLYLQMNSLRVEDTAVYYC | ARGKLLGR      | SNQGT | LVTVSS |
| D            | QVQLVESGGGLVQPGGSLRLSCAAS  | GFTFSSYG | VSWVRQAPGKGLEWVST | ISGSGGST         | YYADSVKGRFTISRDN | SKNTLYLQMNSLRAEDTAVYYC  | AR ...        |       |        |
| E            |                            |          | ... ISGSGGST      | YYPDSVKGRFTISRDN | AKNSLYLQMNSLRDED | TAVYYCA                 | RVRSGGYMDV    | WGQGT | LVTVSS |
| G            | QVQLVESGGGLIQQPGGSLRLSCAAS | EFTVSSNH | MSWVRQAPGKGLEWVSV | LYSGGNT          | YYADSVKGRFTISRDD | SKNTLYLQMNSLRAEDTAVYSC  | ARGKYSSSWYFDL |       |        |
|              | WGRGTLVTVSS                |          |                   |                  |                  |                         |               |       |        |
| H            | QVQLVESGGGLVQPGGSLRLSCAAS  | GFTFSNHG | MNWVRQAPGKGLEWLSY | IGGSGGSI         | RNADSVKGRFTISRDN | AKNSLSLQMNSLRAEDTAMYYC  | ARGGGRLLDY    | WGQGT | LVTVSS |
| I            | QVQLVESGGGLVQPGGSLRLSCAAS  | GFTFSRHA | MSWVRQAPGKGLEWVSA | ISGSGGST         | YYADSVKGRFTISRDN | SKNTLYLQMNSLRAEDMAVYYC  | ARGRGSWYLDY   | WGQGT | LVTVSS |
| J            | QVQLQQSGGGLVQPGGSLRLSCSAS  | GFTFSTYA | IHWVRQAPGKGLEWVSG | ISGTGAST         | YYADSVKGRFTISRDN | AKNSVYLQMNSLRVEDTAVYYC  | ADFGFP        | WGQGT | LVTVSS |

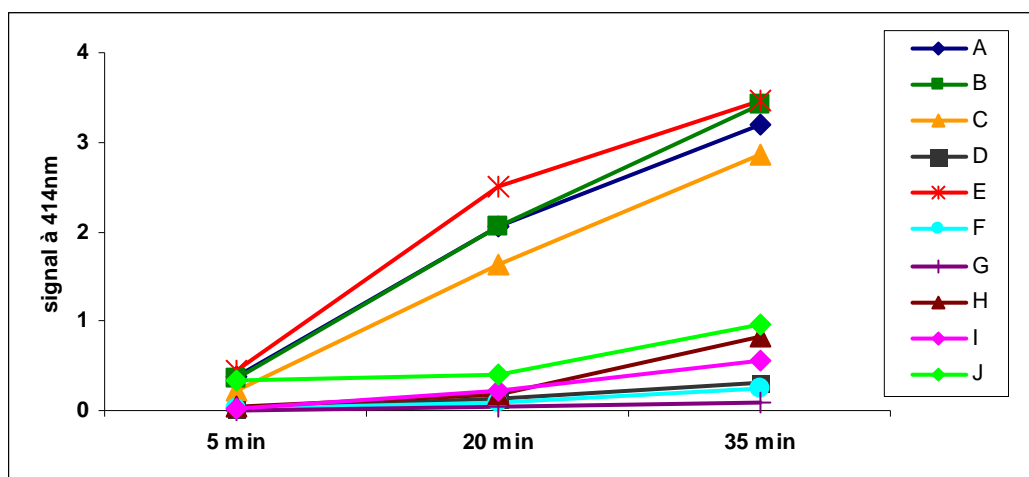
| Light chains |                            |           |                    |                   |                                       |                                      |             |             |  |
|--------------|----------------------------|-----------|--------------------|-------------------|---------------------------------------|--------------------------------------|-------------|-------------|--|
| Profil       | IMGT FR1                   | CDR1      | FR2                | CDR2              | FR3                                   | CDR3                                 | FR4         |             |  |
| A            | SSELTQDPAVSVALGQTVRITCQGD  | SLRSYY    | ASWYQQKPGQAPVLVIY  | GKN               | NRPSGIPDRFSGSSSGNTASLTITGAQAEDEADYYC  | NSRDSSGNPLV                          |             |             |  |
|              | FGGGTKLTVLG                |           |                    |                   |                                       |                                      |             |             |  |
| B            | QSALTQPASVSGSPGQSLTISCTGT  | SSDVGGYKY | VSWEYQQHPGKAPKVMII | DVT               | NRPSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYC | SSYTTSSSTLV                          |             |             |  |
|              | FGGGTKLTVLV                |           |                    |                   |                                       |                                      |             |             |  |
| B'           | QSALTQPASVSGSPGQSLTISCTGT  | SSDVGGYKY | VSWEYQQHPGKAPKVMII | DVT               | NRPSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYC | SSYTTSSSTLV                          |             |             |  |
|              | FGGGTKLTVLG                |           |                    |                   |                                       |                                      |             |             |  |
| C            | EIVLTQSPSSLSASVGDRTITCRAS  | QSISSY    | LNWYQQKPGKAPKLLII  | GAS               | SLQSVVPSRFGSGSGTEFTLTVSGLQPEDFATYFC   | QQAHSFPPT                            |             |             |  |
|              | FGGGTKLEIKR                |           |                    |                   |                                       |                                      |             |             |  |
| D            |                            |           | ...                | LAWYQQKPGQPPKLLII | WAX                                   | TRESGVPDRFSGSGSGTDFTLTIRSLQPEDFATYYC | QQANSEPLT   | FGGGTKVDIKR |  |
| E            | DIVMTQSPSTLSASIGDRVTITCRAS | EGIYHW    | LAWYQQKPGKAPKLLII  | KAS               | SLASGAPLRYSGSGSGTDFTLTISLQPDFFATYYC   | QQYHTISRT                            | FGPGTKVXXLG |             |  |
| G            | SSELTQDPAVSVALGQTVRITCQGD  | SLRSYY    | ASWYQQKPGQAPVLVIY  | GKN               | NRPSGIPDRFSGSSSGNTASLTITGAQAEDEADYYC  | NSRDSTNNQL                           | FGGGTKVTVLG |             |  |
| H            | QSALTQPASVSGSPGQSLTISCTGT  | SSDVGGYKY | VSWEYQQHPGKAPKVMII | DVT               | NRPSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYC | SSYTTSSSTLV                          |             |             |  |
|              | FGGGTKLTVLG                |           |                    |                   |                                       |                                      |             |             |  |
| I            | QSALTQPPASGSPGQSVTISCTGT   | SSDVGGYNY | VSWEYQQHPGKAPKLMII | EVS               | KRPSGVSNRFGSGSKSGNTASLTISGLQAEDEADYYC | SSYTTSSSWV                           | FGGGTKVTVLS |             |  |
| J            | SSELTQDPAVSVALGQTVRITCQGD  | SLRSYY    | ASWYQQKPGQAPVLVIY  | GKN               | NRPSGIPDRFSGSSSGNTDSLITITGAQAEDEADYYC | NSRDSSGKHVV                          | FGGGTKLTVLG |             |  |

**Figure 22 : Séquences des scFv correspondant aux différents profils obtenus par fingerprinting.**

La distribution des acides aminés au sein des chaînes lourde et légère a été identifiée grâce à l'outil IMGT/V-QUEST (Brochet et al., 2008).



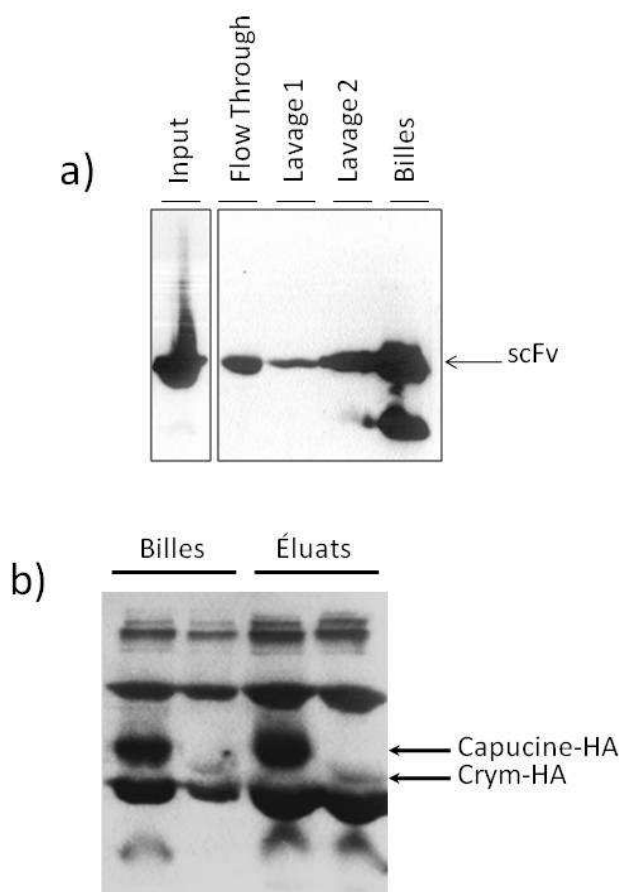
Grâce au tag Myc qui est fusionné aux scFv, nous avons pu analyser par ELISA les affinités des scFv des différents profils pour le fragment recombinant de Capucine (comme décrit par Schier et al., 1995). Les scFv des profils de *fingerprinting* A, B, C et E s'avèrent avoir les meilleures affinités pour le fragment de Capucine (Figure 23).



**Figure 23 : Test d'affinité des scFv pour un fragment recombinant de Capucine par ELISA.**

A à J : Dix parmi les onze scFv différents.

En raison de son affinité pour la Capucine, un scFv du profil de *fingerprinting* « A » a été sélectionné. Nous avons sous-cloné l'ADNc codant ce scFv dans un plasmide d'expression ajoutant une étiquette 6-histidine au scFv, afin de faciliter sa purification (comme décrit par Schier et al., 1995). Les scFv ont été produits et purifiés sur colonne de nickel (comme décrit par Schier et al., 1995 et Nielsen et al., 2006). Le succès de la purification a été vérifié par Western blot en utilisant des anticorps anti-Myc (Figure 24A). Puis les billes sur lesquelles les scFv ont été retenues grâce à leur étiquette 6-histidine ont été utilisées pour immunoprécipiter la Capucine fusionnée au tag HA produite *in vitro* dans des lysats de réticulocytes (Figure 24B). Le résultat étant positif nous avons voulu tester si les scFv élués des billes pouvaient reconnaître la Capucine endogène dans des extraits protéiques totaux de striatum de souris. Une seule expérience de Western blot a pour le moment été réalisée et aucun signal n'a été observé.



**Figure 24 : Production d'un scFv du profil « A ».**

a) Purification du scFv. Pistes 1 : input ; 2 : flow through ; 3 : lavage NaCl 1M ; 4 : lavage imidazole 35mM, 5 : billes de nickel. Les signaux sont révélés par un anticorps anti-Myc.

b) Immunoprécipitation de la Capucine étiquetée par HA produite *in vitro*. Pistes 1 et 3 immunoprécipitation de la Capucine-HA ; 2 et 4 immunoprécipitation de la Crym-HA (Crystaline). Les signaux sont révélés par un anticorps anti-HA. La protéine Crym-HA, produite dans le même système sert de témoin. Plusieurs signaux non spécifiques sont également observés.

Grâce à la technique de « *phage display* », nous avons obtenu, produit et purifié un anticorps, dont le titre reste à mesurer. Cet anticorps permet d'immunoprécipiter la Capucine produite *in vitro*. Une première analyse en Western blot utilisant ce scFv comme anticorps primaire n'a pas donné de résultat positif. Mais une mise au point du Western blot serait nécessaire. On pourrait en premier lieu tester différentes concentrations de scFv. En outre, il faudrait affiner les conditions d'extraction des protéines des régions du cerveau. En effet, la Capucine est une protéine transmembranaire et il est possible que de ce fait elle soit particulièrement difficile à extraire. Enfin, il reste encore plusieurs scFv (au moins trois profils de *fingerprinting*) qui paraissent intéressants d'après les tests d'affinité effectués par ELISA. Leur production et leur purification pourraient être envisagées.

## 5. Discussion

Dans cette étude le rôle biologique de la Capucine a été exploré à l'aide d'une lignée de souris knock-out pour ce gène construite au laboratoire. Ces souris ne présentent pas d'anomalies particulières, elles sont apparemment en bonne santé et fertiles. L'absence de transcrits codant la Capucine dans le cerveau de ces animaux a été démontrée par Northern blot. Nous avons également vérifié par RT-PCR quantitative que le knock-out de Capucine n'est pas compensé par la surexpression de son paralogue le plus proche, SynDIG1, dans le striatum. Cependant, une analyse transcriptomique globale pourrait être envisagée afin d'identifier les gènes dont l'expression est modifiée chez les animaux knock-out. Ceci pourrait nous donner des indications sur la fonction de Capucine.

Etant donné que la Huntingtine est exprimée de façon ubiquitaire, les raisons de la vulnérabilité particulière des neurones striataux sont encore mal comprises. L'expression de Capucine est enrichie dans le striatum et perturbée dans des modèles de la maladie de Huntington (de Chaldée et al., 2006). Le rôle potentiel de la Capucine dans le développement de cette maladie a donc été exploré dans un modèle lentiviral de la maladie de Huntington. Nos résultats montrent que ni le knock-out, ni l'augmentation intrastriatale de l'expression de Capucine ne conduisent à une modification de la toxicité d'un fragment de la Huntingtine mutée. L'effet potentiel de la Capucine pourrait être étudié dans d'autres modèles de la maladie de Huntington, en particulier des modèles génétiques, où la mutation est présente dès le début du développement. Cette approche est envisageable car les souris knock-out pour Capucine sont fertiles. Dans la mesure où le gène de la Capucine est exprimé à des niveaux élevés et de manière spécifique dans le striatum, il serait également intéressant d'analyser le comportement locomoteur des animaux knock-out (par exemple test du rotarod et l'activité spontanée en openfield). Dans cette perspective, un back-cross avec la souche C57BL6/J est en cours.

La fonction de la Capucine n'est cependant toujours pas connue. Son paralogue SynDIG1 a été décrit récemment (Kalashnikova et al., 2010). La connaissance de son rôle nous donne une nouvelle piste à explorer. Il s'agit d'une protéine transmembranaire dont l'extrémité C-terminale est extracellulaire et l'extrémité N-terminale cytoplasmique, ce qui indique qu'une seule des deux régions hydrophobes est transmembranaire. Contrairement à la Capucine, SynDIG1 n'est pas présent dans l'appareil de Golgi, mais dans les endosomes. Son expression a été localisée dans les cellules de Purkinje du cervelet et dans l'hippocampe.

SynDIG1 joue un rôle critique dans le développement des synapses contenant les récepteurs glutamatergiques de type AMPA. En outre les auteurs ont observé un pic d'expression de SynDIG1 deux semaines après la naissance chez le rat, ce qui correspond à la période de synaptogénèse. L'expression de la Capucine a été observée dans le striatum à partir de cinq jours après la naissance. Le rôle potentiel de la Capucine dans le développement des synapses pourrait de ce fait être examiné. La production d'anticorps que j'ai initiée devrait permettre d'évaluer dans un premier temps si la Capucine est localisée au niveau synaptique et d'identifier ces synapses. L'expression de la Capucine par les neurones de projection du striatum (à la fois striatopallidaux et striatonigraux) a été démontrée (Heiman et al., 2008). Par conséquent, si la Capucine est détectée dans le striatum, elle pourrait être localisée dans l'élément postsynaptique des synapses glutamatergiques (afférences corticales et thalamiques) ou dopaminergiques (afférences nigrales). Si la Capucine est détectée dans les structures de projection du striatum (pallidum et substance noire), elle pourrait être localisée dans l'élément présynaptique des synapses GABAergiques. Ces hypothèses pourront être vérifiées par des expériences de colocalisation avec des marqueurs de ces synapses (par exemple SAP102 pour l'élément postsynaptique des synapses glutamatergiques, les récepteurs D1 et D2 pour l'élément postsynaptique des synapses dopaminergiques et le transporteur du GABA pour l'élément présynaptique des synapses GABAergiques). Par ailleurs, les anticorps contre la Capucine permettront d'identifier ses partenaires par des expériences de co-immunoprécipitation.

## IV. Etude fonctionnelle d'AGPAT4

Une étude comparative de l'expression des gènes dans plus de dix régions du cerveau de souris a été effectuée en utilisant la technique SAGE avant mon arrivée en thèse. Des gènes principalement exprimés dans chacune de ces régions ont été identifiés (Brochier et al., 2008). Ces travaux ont montré que l'expression d'Agpat4 est enrichie dans l'aire tegmentaire ventrale et la substance noire compacte. Les neurones dopaminergiques de la substance noire étant sélectivement atteints lors de la maladie de Parkinson, Agpat4, dont la fonction est inconnue, pourrait se montrer intéressant quant à sa participation éventuelle au développement de cette maladie neurodégénérative ou pour la compréhension des fonctions spécifiques de ces régions du cerveau.

D'après des homologues de séquences, Agpat4 a été classée dans une famille d'acyltransférases de lysophospholipides (Lewin et al., 1999). Il s'agit d'enzymes impliquées dans le métabolisme des glycérophospholipides, dont certaines sont aujourd'hui déjà bien caractérisées.

### 1. Métabolisme des glycérophospholipides

Les glycérophospholipides sont des molécules amphipathiques (ou amphiphiles) car elles présentent à la fois une partie hydrophobe due aux acides gras (Tableau 1) et une partie hydrophile due à la présence d'une tête polaire formée par l'ester phosphorique. L'élément de base des glycérophospholipides est l'acide phosphatidique qui consiste en un squelette glycérol, estérifié par deux acides gras en positions *sn1* et *sn2*, et par un acide phosphorique en position *sn3*. L'acide gras en position *sn1* est en général court et saturé et celui en position *sn2* long et insaturé. L'acide phosphorique peut être estérifié par un groupement qui détermine la classe du glycérophospholipide. Il peut s'agir de la choline, de l'éthanolamine, de l'inositol, de la sérine ou d'un autre glycérol formant ainsi une phosphatidylcholine (PC), une phosphatidyléthanolamine (PE), un phosphatidylinositol (PI), une phosphatidylsérine (PS), un phosphatidylglycérol (PG) ou encore une cardiolipine (CL), les phospholipides majeurs chez les eucaryotes (Figure 25). Les glycérophospholipides ne portant qu'une seule chaîne d'acide gras sont désignés par le préfixe lyso-.

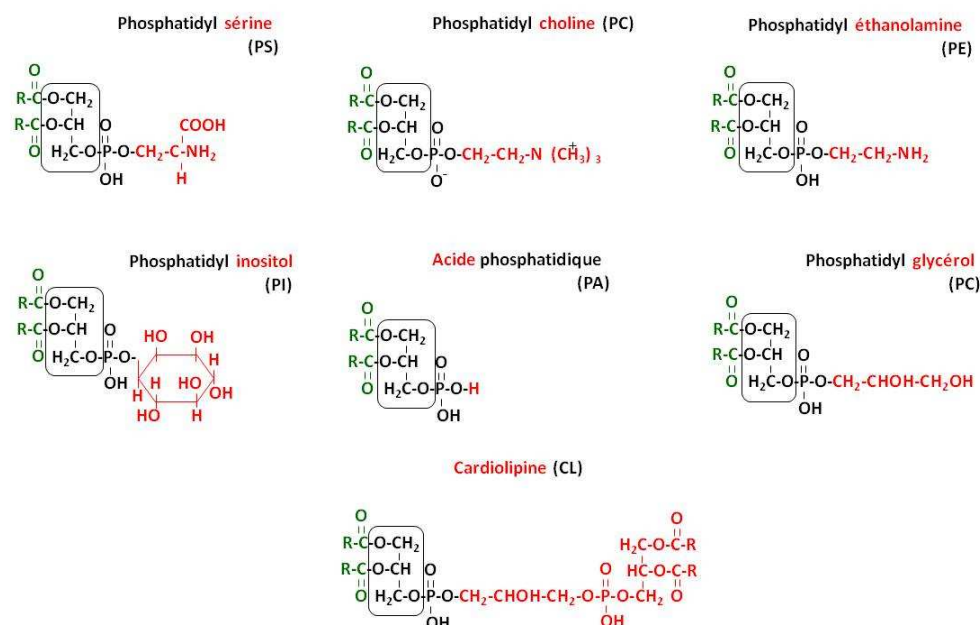


Figure 25 : Structure des glycérophospholipides.

| Symbole            | Structure chimique   | Non systématique de l'acide                 | Nom commun de l'acide |
|--------------------|--|---|-----------------------|
| C12 : 0            | CH <sub>3</sub> (CH <sub>2</sub> ) <sub>10</sub> COOH  | Dodécanoïque                                | laurique              |
| C14 : 0            | CH <sub>3</sub> (CH <sub>2</sub> ) <sub>12</sub> COOH  | Tétradécanoïque                             | myristique            |
| C16 : 0            | CH <sub>3</sub> (CH <sub>2</sub> ) <sub>14</sub> COOH  | Hexadécanoïque                              | palmitique            |
| C18 : 0            | CH <sub>3</sub> (CH <sub>2</sub> ) <sub>16</sub> COOH  | Octadécanoïque                              | stéarique             |
| C20 : 0            | CH <sub>3</sub> (CH <sub>2</sub> ) <sub>18</sub> COOH  | Icosanoïque                                 | arachidique           |
| C22 : 0            | CH <sub>3</sub> (CH <sub>2</sub> ) <sub>20</sub> COOH  | Docosanoïque                                | béhénique             |
| C24 : 0            | CH <sub>3</sub> (CH <sub>2</sub> ) <sub>22</sub> COOH  | Tétracosanoïque                             | lignocérique          |
| C18 : 1(9)         | CH <sub>3</sub> (CH <sub>2</sub> ) <sub>7</sub> CH=CH(CH <sub>2</sub> ) <sub>7</sub> COOH                                      | cis-9-octadécénoïque                        | oléique               |
| C18 : 2(9,12)      | CH <sub>3</sub> (CH <sub>2</sub> ) <sub>4</sub> CH=CHCH <sub>2</sub> CH=CH(CH <sub>2</sub> ) <sub>7</sub> COOH                 | cis,cis-9,12-octadécadiénoïque              | linoléique            |
| C20 : 4(5,8,11,14) | CH <sub>3</sub> (CH <sub>2</sub> ) <sub>4</sub> (CH=CHCH <sub>2</sub> ) <sub>3</sub> CH=CH(CH <sub>2</sub> ) <sub>3</sub> COOH | cis,cis,cis,cis-5,8,11,14-icosatétraénoïque | arachidonique         |

Tableau 1: Les principales chaînes d'acides gras

## 1.1. Fonction des glycérophospholipides

Les glycérophospholipides sont les principaux lipides de structure des membranes eucaryotes. La phosphatidylcholine (PC) représente plus de 50% des phospholipides membranaires (Van Meer, 2005). Elle s'auto-organise spontanément pour former une bicouche lipidique dans laquelle chaque PC a une géométrie moléculaire presque cylindrique, avec les queues d'acide gras se faisant face et les têtes polaires orientées vers la phase aqueuse du milieu extérieur. La phosphatidyléthanolamine (PE) prend une forme moléculaire conique en raison de la taille relativement petite de sa tête polaire. L'inclusion de PE dans les

bicouches de PC crée une tension sur la membrane, ce qui peut être utilisé par exemple pour le bourgeonnement, la fission et la fusion (Marsh, 2007). Les lipides ne formant pas de bicouches, comme la PE et la cardiolipine (CL), prennent en charge les protéines membranaires et peuvent moduler leurs activités (Bogdanov et al., 2008).

Les phospholipides sont non seulement les principaux constituants des membranes biologiques, mais ils participent également au contrôle de diverses fonctions cellulaires, telles que la régulation de la transcription (Boeneman and Crooke, 2005), de la transduction du signal (Wang et al., 2006), ou du repliement des protéines (Bogdanov et al., 1999). Par exemple, l'acide lysophosphatidique (LPA) et l'acide phosphatidique (PA) jouent un rôle important dans différentes réponses cellulaires de signalisation. Le LPA est connu comme facteur de croissance lipidique impliqué dans la régulation de nombreuses réponses cellulaires par l'activation de récepteurs couplés aux protéines G (Li et al., 2003). Le PA est connu comme second messenger lipidique participant à des événements divers de signalisation intracellulaire notamment la mobilisation du calcium intracellulaire et l'entrée du calcium extracellulaire (Salmon et Honeyman, 1980; Putney et al., 1980), et régule différentes protéines de signalisation, y compris plusieurs protéines kinases et phosphatases (English et al., 1996). L'acide phosphatidique est aussi un médiateur de certaines voies de signalisation de la cellule associées à l'oncogenèse. Il s'agit notamment de la voie Ras / Raf / ERK et Akt / mTOR (Bonham et al., 2003). mTOR (*mammalian target of rapamycin*) appartient à la famille des phosphatidylinositol-3 kinases et est impliqué dans l'initiation de la traduction chez les mammifères. mTOR régule la croissance et la prolifération cellulaire via ses cibles telles que la kinase ribosomique S6K1 et la protéine de liaison au facteur d'initiation de la traduction E4-BP1 (Chen, 2004). Le PA a été identifié comme un activateur mitogénique de la voie de signalisation de mTOR. Il se lie à la protéine mTOR pour initier la synthèse des protéines et génère des signaux de survie dans les cellules tumorales du sein (Chen, 2004). Il intervient dans la translocation de la sérine/thréonine kinase Raf-1 du cytosol vers la membrane où elle est activée par le régulateur membranaire Ras, ce qui va déclencher l'activation de la cascade de signalisation de MAPK (*mitogen-activated protein kinase*) (Rizzo et al., 2000). Le PA a aussi des fonctions anti-apoptotiques (Deng et al., 2003).

## 1.2. Biosynthèse des glycérophospholipides

Les glycérophospholipides PC, PE, PS et PI sont synthétisés sur la surface cytosolique du réticulum endoplasmique et de l'appareil de Golgi. Le PG et la CL sont synthétisés dans les mitochondries. La PE peut également être générée par la décarboxylation de la PS dans les mitochondries (Henneberry et al., 2002). Deux voies de biosynthèse sont employées par la cellule :

- La Voie de Kennedy (synthèse *de novo*)
- Le Cycle de Land (voie de recyclage)

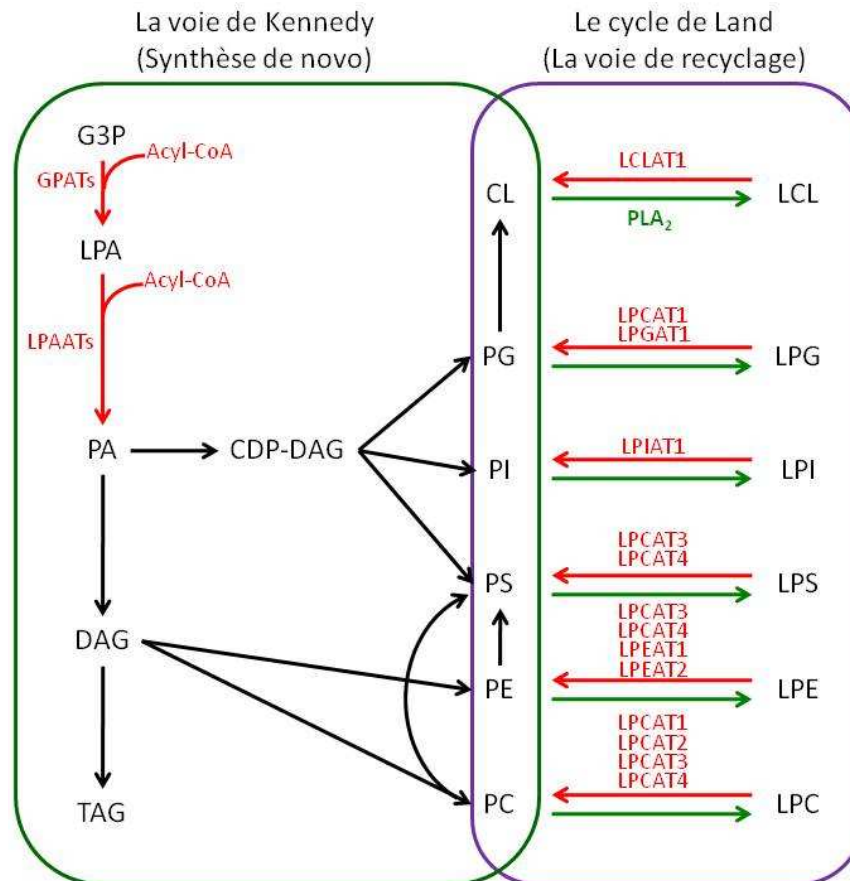
### 1.2.1. La voie de Kennedy

La synthèse complète *de novo* des différents glycérophospholipides à partir du glycérol a été décrite par Kennedy et Weiss il y a une cinquantaine d'année (Kennedy and Weiss, 1956). La première étape consiste en l'acylation du glycérol-3-phosphate en position *sn1* (Figure 26). La réaction est catalysée par une glycérol-*sn*-3-phosphate acyltransférase (GPAT) qui utilise un acyl-CoA pour produire un acide lysophosphatidique (LPA) ou 1-acyl-*sn*-glycérol-3-phosphate (Kent, 1995). Le LPA est ensuite converti en acide phosphatidique (PA) par une acylation en position *sn2*. Cette réaction est catalysée par une 1-acylglycérol-*sn*-3-phosphate acyltransférase (AGPAT) également appelé *lysophosphatidic acid acyltransferase* (LPAAT). L'acide phosphatidique formé suite à cette réaction occupe un point central dans les voies de biosynthèse des lipides. Il peut être converti en cytidine diphospho diacylglycérol (CDP-DAG), le précurseur de la biosynthèse de la PS, du PI, et du PG, lui-même précurseur de la synthèse de CL. Alternativement l'acide phosphatidique peut être déphosphorylé en produisant un diacylglycérol, le précurseur du triacylglycérol (TAG), de la PC, et de la PE. La PE peut aussi être convertie en une PS (Coleman and Lee, 2004).

### 1.2.2. Le Cycle de Land

Les glycérophospholipides sont d'abord synthétisés dans la voie *de novo* et ensuite la composition des acides gras situés en position *sn2* peut être modifiée dans la voie de recyclage par l'action combinée de la phospholipase A<sub>2</sub> et de différentes lysophospholipide acyltransférases (LPLAT) (Land, 1958). Il s'agit d'une réaction bidirectionnelle où des lysoglycérophospholipides sont convertis en glycérophospholipides et *vice versa* (Figure 26).





**Figure 26 : Les voies de biosynthèse des glycérophospholipide**

Les glycérophospholipides sont d'abord synthétisés par la voie de synthèse *de novo* (La voie de Kennedy) et ensuite modifiés par la voie de recyclage (Le cycle de Land). Les flèches rouges indiquent les réactions catalysées par les acyltransférases et les vertes celles réalisées par les PLA<sub>2</sub> (d'après Shindou et Shimizu, 2009).

### 1.3. Les membres de la famille Agpat

L'alignement des séquences d'acides aminés entre différentes acyltransférases du métabolisme des glycérophospholipides a révélé quatre régions (motifs d'acyltransférases I-IV : NH-(X)<sub>4</sub>-D (motif I), GVIFIDR (motif II), EGTR (motif III), et IVPIVM (motif IV)) bien conservées entre diverses espèces hormis les plantes (Tableau 2) (Lewin et al., 1999). Plusieurs acides aminés invariables ont été identifiés dans ces motifs. Ils ont un rôle potentiel dans la fonction catalytique de l'enzyme (dans les motifs I, III et IV) ou dans la liaison du substrat (dans les motifs II et III). Cette hypothèse a été appuyée par les résultats des études de mutagenèse dirigée sur GPAT chez *E. coli* (Lewin et al., 1999) et sur la GPAT1 de souris (Dircks et al., 1999).

|                            | Block I        | Block II | Block III     | Block IV | Accession |
|----------------------------|----------------|----------|---------------|----------|-----------|
| <b>GPAT</b>                | 303            | 348      | 382           | 417      |           |
| E. coli                    | VPCHRSHMDYLL   | GAFFIRR  | YFVEGGRSRTGR  | ITLIPIYI | 130326    |
| H. influenza               | VPCHRSHIDYLL   | GAFFIRR  | YFIEGGRSRTGR  | ISIVPVYV | 1172533   |
| M. tuberculosis            | AFSHRSYLDGML   | GAIFIRR  | WSIEGGRTRTGK  | VYLVPTSI | 2791522   |
| C. elegans                 | ICLHRSHLDILSM  | NTIFIRR  | FFLEGTRSRFGK  | ISIIPVVF | 1458332   |
| M. musculus                | LPVHRSHIDYLL   | GGFFIRR  | IFLEGTRSRSGK  | ILVIPVGI | 2498786   |
| R. norvegicus              | LPVHRSHIDYLL   | GGFFIRR  | IFLEGTRSRSGK  | ILVIPVGI | 2444459   |
| <b>LPAAT</b>               |                |          |               |          |           |
| E. coli                    | IANHQNNYDMVTA  | GNLLDR   | MFPEGTRSR .GR | VPIIPVCV | 1789395   |
| S. typhimurium             | IANHQNNYDMVTA  | GNLLDR   | MFPEGTRSR .GR | VPIIPVCV | 154236    |
| H. influenzae              | IGNHQNNYDMVTI  | GNIFLDR  | MFPEGTRNR .GR | VPIIPVVC | 1573737   |
| M. genitalium              | VANHKSNLDPLVL  | DCVFDR   | VFAEGTRIL .SN | VPILPVS  | 1045898   |
| S. cerevisiae              | IANHQSTLDIFML  | GTYFLDR  | VFPEGTRSTSEL  | IPIVPVV  | 464422    |
| C. elegans                 | ICNHQSSLDILSM  | NTIFDR   | VFPEGTRNREGG  | IPIPVVF  | 1403001   |
| M. musculus                | VSNHQSSLDLLGM  | GIIFDR   | VFPEGTRNHNGS  | VPIIPIVF | 2467310   |
| H. sapiens alpha           | VSNHQSSLDLLGM  | GVIFDR   | VFPEGTRNHNGS  | IPIPVVF  | 2155238   |
| H. sapiens beta            | VSNHQSSLDMMGL  | GVFFINR  | IYPEGTRNDNGD  | VPIIPVVY | 2155240   |
| <b>DHAPAT (H. sapiens)</b> | LPSHRSYIDFLML  | GAFFMRR  | FFLEGTRSRSAK  | TYLVPISI | 3258645   |
| <b>LPEAT (E. coli)</b>     | TPNHVSFIDGILL  |          | IFPEGTRITTTG  | ATVIPVRI | 290403    |
| <b>GPAT consensus</b>      | φ--HRS-φDφφ--  | G-φFIRR  | φFφEGTRSR-GK  | φφφφPφφφ |           |
| <b>LPAAT consensus</b>     | φ-NHQSS-φDφφ-- | G-φFIDR  | φFPEGTR---G-  | φPφφPφφφ |           |

**Tableau 2 : Les motifs conservés dans les différentes acyltransférases**  
Tableau réalisé en utilisant l'algorithme ClustaWL (D'après Lewin et al., 1999)

### 1.3.1. Fonctions des différents membres de la famille Agpat

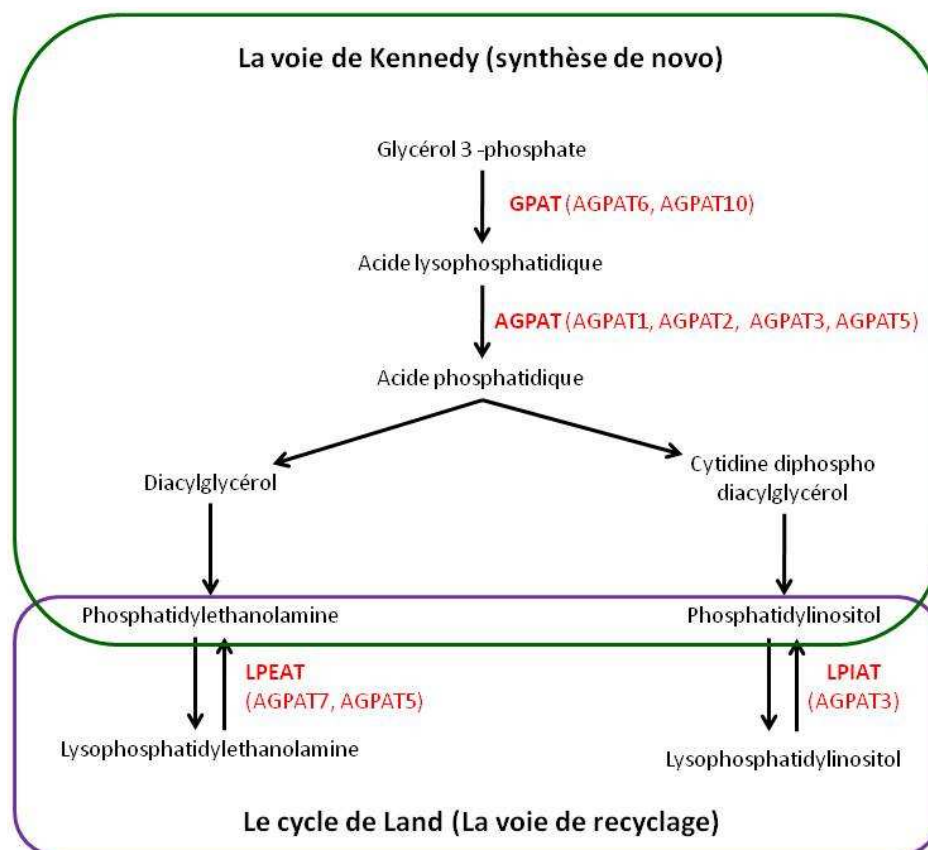
Les membres de cette famille peuvent être classés en fonction de leur activité dans les groupes ci-dessous (Shindou et Shimizu, 2008 ; Hishikawa et al., 2008). Plusieurs de ces enzymes, initialement appelées Agpat en raison de leurs homologies de séquence, se sont avérées catalyser d'autres réactions (Figure 27).

#### Enzymes de la voie de Kennedy :

- GPAT (*glycérol-3-phosphate acyltransférases*) : Gpat1  
Gpat2  
Gpat3 (= Agpat10)  
Gpat4 (= Agpat6)
- LPAAT (*acide lysophosphatidique acyltransférases*) : Agpat1  
Agpat2  
Agpat3 (également LPIAT)  
Agpat5 (également LPEAT)

#### Enzymes du cycle de Land :

- LPLAT (*lysophospholipid acyltransférases*) :  
LPEAT (*lysophatidyléthanolamine acyltransférases*) : Agpat7 (= Lpcat4 = Ayt13)  
LCLAT (*lysocardiolipine acyltransférases*) : Agpat8 (= Lycat = Alcat = Lclat1)  
LPCAT (*lysophosphatidylcholine acyltransférase*) : Agpat9 (= Lpcat1 = Ayt12)  
Agpat11 (= Lpcat2 = Ayt11 = LysoPAFAT)
- LPGAT (*lysophosphatidylglycérol acyltransférase*) : Lpgat1
- Activité inconnue : Lpcat2b (= Ayt16 = AT like 1B)  
Agpat4



**Figure 27 : Les fonctions de différents membres de la famille AGPAT dans la biosynthèse des glycérophospholipides**

Les différentes enzymes nommées AGPATs interviennent à différents niveaux dans la synthèse des glycérophospholipides. Ce schéma montre la redondance des AGPATs, à l'exemple d'AGPAT3 et AGPAT5.

#### 1.3.1.1. Membres les mieux caractérisés

Agpat1 et Agpat2 sont les mieux caractérisées et leur activité AGPAT est bien décrite. AGPAT1 possède quatre régions hydrophobes qui pourraient correspondre à des domaines transmembranaires potentiels (Aguado and Campbell, 1998 ; Leung 2001). Une topologie détaillée et des études de mutagenèse suggèrent un modèle avec quatre domaines transmembranaires avec le motif acyltransférase I localisé dans le cytosol et le motif III situé dans le lumen du réticulum endoplasmique. Les motifs I et IV sont impliqués dans la liaison de l'acyl-CoA et dans la catalyse, tandis que les motifs II et III sont impliqués dans la liaison du LPA. La localisation des motifs II et III dans une région hydrophobe peut être nécessaire pour la liaison des accepteurs d'acyle hydrophobes (Yamashita et al., 2007). L'activité AGPAT et la localisation d'AGPAT1 dans le réticulum endoplasmique ont été confirmés par surexpression dans des cellules de mammifère et d'insecte (Aguado and Campbell,

1998 ; Kume et al., 1997 ; Stamps et al., 1997 ; West et al., 1997). AGPAT1 utilise de préférence certains acides gras (C12-16:0, C16:1, C18:2 et C18:3, suivis par C18:0, C18:1 et C20:4) et a une faible activité avec d'autres comme C20:0 et C24:0 (Aguado and Campbell, 1998 ; Hollenback et al., 2006). En outre AGPAT1 peut également catalyser la synthèse de LPA à partir de PA, une réaction inverse de l'activité AGPAT (Yamashita et al., 2007). Cette activité réversible pourrait être impliquée dans la régulation des niveaux de LPA et de PA qui peuvent agir comme des molécules de signalisation. Les études de surexpression d'AGPAT1 dans les adipocytes 3T3-L1 et les myotubes C2C12 suggèrent son rôle potentiel dans la gestion de l'énergie dans le tissu adipeux et le muscle (Ruan and Pownall, 2001).

AGPAT2 est le seul membre connu de la famille dont les mutations soient responsables d'une maladie humaine, la lipodystrophie généralisée congénitale (Agarwal et al., 2002). Des souris déficientes pour AGPAT2 ont également été développées (Cortes et al., 2009), ce qui fait d'AGPAT2 le membre de la famille Agpat dont la physiologie est le mieux étudiée. La séquence d'acides aminés d'AGPAT2 humaine montre 35% d'identité avec AGPAT1 (West et al., 1997). Le profil d'hydrophobicité d'AGPAT2 est similaire à celui d'AGPAT1, ce qui suggère qu'il possède des domaines transmembranaires d'ancrage dans la membrane du réticulum endoplasmique (Eberhardt et al., 1997 ; Leung, 2001 ; Yamashita et al., 2007). La localisation subcellulaire d'AGPAT2 dans le réticulum endoplasmique a été vérifiée (Eberhardt et al., 1999 ; Gale et al., 2006). L'activité enzymatique d'AGPAT2 a été confirmée par surexpression chez *E. coli*, dans des cellules d'insecte et de mammifère (Eberhardt et al., 1999 ; Hollenback et al., 2006 ; Lu et al., 2005 ; West et al., 1997). Les préférences de substrat d'AGPAT2 ont été étudiées en détail et il a été montré qu'AGPAT2 possède une spécificité de substrat stricte pour les LPA (Hollenback et al., 2006). Son substrat préféré est le 1-oléoyl-LPA. Plusieurs acides gras peuvent être incorporés, y compris les acides gras C14:0, C16:0, C18:1 et C18:2, et une faible incorporation de C18:0 et C20:4 a été observée.

Bien que d'autres enzymes AGPAT soient également exprimées dans le tissu adipeux, un rôle essentiel d'AGPAT2 dans la synthèse du TAG dans les adipocytes humains a été établi à partir d'études de patients atteints de lipodystrophie généralisée congénitale (CGL) et présentant des mutations dans le gène AGPAT2 (Agarwal et al., 2002). La CGL est une maladie autosomique récessive caractérisée par une absence de tissu adipeux depuis la naissance, une résistance sévère à l'insuline, une hypertriglycémie, une stéatose hépatique, et un diabète précoce (Agarwal and Garg, 2003 ; Hegele et al., 2007). La plupart des mutations d'AGPAT2 causant la CGL entraînent une perte quasi-complète de l'activité

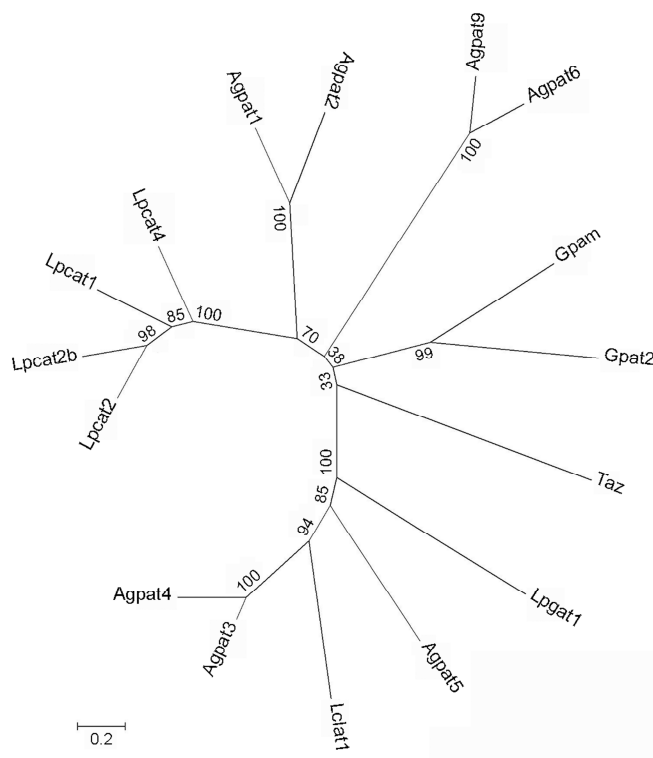
d'AGPAT2 *in vitro* (Haque et al., 2005). La plupart des tissus adipeux viscéraux et la graisse sous-cutanée manquent chez les patients souffrant de CGL, mais les tissus adipeux mécaniques tels que la graisse sous-cutanée de la paume des mains, des plantes de pied et du cuir chevelu ainsi que la graisse des régions péri-articulaires et rétro-orbitaire sont maintenues (Simha and Garg, 2003). Ceci suggère que d'autres isoformes d'AGPAT pourraient être actives dans ces tissus adipeux mécaniques.

Le knock-down d'AGPAT2 réprime également l'expression d'autres gènes AGPAT (AGPAT1, 3, 4 et 5) et augmente les niveaux de plusieurs espèces de phospholipides au cours de la différenciation des adipocytes. Ces résultats suggèrent que ces AGPAT modulent les taux de phospholipides cellulaires, qui peuvent influencer sur la différenciation adipocytaire (Gale et al., 2006). Les souris knock-out pour AGPAT2 ont un phénotype qui n'est pas compensé par les autres AGPAT, ce qui indique qu'AGPAT2 a un rôle physiologique non redondant. Elles présentent un défaut de tissu adipeux sous-cutané et intra-abdominal blanc et de tissu adipeux brun interscapulaire. Chez ces souris, la graisse représente seulement 2% de la masse corporelle contre 25% chez les souris sauvages. 80% des souriceaux déficients pour AGPAT2 meurent à l'âge de 3 semaines, probablement d'hyperglycémie et/ou d'hypothermie (Cortes et al., 2009).

Une surexpression d'AGPAT2 a été observée dans certains cancers humains (leucémie, lymphome non Hodgkinien, myélomes multiples, cancer de l'ovaire, cancer de l'endomètre), et son inhibition spécifique a été étudiée comme thérapie potentielle. Les inhibiteurs d'AGPAT2 induisent un arrêt de croissance, l'apoptose ou la nécrose dans diverses lignées cellulaires tumorales (Coon et al., 2003 ; Douvas et al., 2006 ; Hideshima et al., 2005 ; La Rosée et al., 2006 ; Pagel et al., 2005 ; Springett et al., 2005). L'activité antitumorale des inhibiteurs d'AGPAT2 a été démontrée chez des souris porteuses de tumeurs (Pagel et al., 2005 ; Springett et al., 2005). Il a été montré chez le xénope que les inhibiteurs d'AGPAT2 bloquent les voies Ras/Raf/ERK et inositol-3-phosphate kinase/Akt, ce qui suggère que le PA produit par AGPAT2 est impliqué dans ces voies (Coon et al., 2003).

### 1.3.1.2. Membres présentant les homologies les plus fortes avec Agpat4

Agpat3, Agpat8 (Lclat1) et Agpat5 sont les homologues les plus proches d'Agpat4 (Figure 28).



**Figure 28 : Arbre phylogénétique de différents membres de la famille AGPAT chez la souris**

Cet arbre a été généré à partir des séquences protéiques de tous les membres connus de la famille Agpat par la méthode « Neighbor-Joining ». Les séquences ont été alignées par le logiciel ClustalW. Cent « bootstraps » ont été réalisés.

Agpat3 est présente dans le réticulum endoplasmique, l'enveloppe nucléaire et l'appareil de Golgi (Prasad et al., 2011 ; Schmidt et Brown, 2009). La topologie d'Agpat3 est semblable à celle d'Agpat1. Agpat3 possède deux domaines transmembranaires. L'un d'eux sépare les motifs I et II, qui constituent des domaines fonctionnels essentiels pour l'activité enzymatique. Le motif I (domaine catalytique putatif) est situé dans le cytoplasme et le motif II (domaine de liaison putatif) est dans le lumen du réticulum endoplasmique et de l'appareil de Golgi (Schmidt et al., 2010). D'après Schmidt et al., (2010) le motif II recouvre partiellement le premier domaine transmembranaire et pourrait donc lier des substrats dans la membrane et les canaliser vers le motif I. Agpat3 pourrait également avoir des activités des deux côtés de la membrane, comme le transfert des chaînes acyle d'un feuillet à l'autre. Le substrat préféré d'Agpat3 est le 1-oleoyl-LPA et son donneur d'acyle préféré est l'oléoyl-

CoA. Néanmoins elle peut utiliser d'autres substrats comme la LPC, la LPS et le LPI avec l'arachidonoyl-CoA comme donneur d'acyl (Prasad et al., 2011 ; Yuki et al., 2009), indiquant son double rôle dans la voie *de novo* et le recyclage de synthèse des glycérophospholipides.

AGPAT5 est présente dans le réticulum endoplasmique, l'enveloppe nucléaire et les mitochondries (Prasad et al., 2011). Sa préférence de substrats est similaire à celle d'AGPAT3, mais elle peut également acyler la LPE. Dans ce cas, elle incorpore C18:1 préférentiellement à C20:4.

AGPAT8 a seulement 20% de l'activité spécifique d'AGPAT2. En revanche il a été démontré qu'AGPAT8 possède une activité acyl-CoA:lysocardiolipin acyltransférase (ALCAT) (Cao et al., 2004 ; Agarwal et al., 2006). En réponse à un stress oxydant, AGPAT8 catalyse le remodelage des cardiolipines, ce qui conduit à un dysfonctionnement mitochondrial et à une résistance à l'insuline (Li et al., 2010). AGPAT8 catalyse également l'acylation de précurseurs de la cardiolipine (Cao et al., 2009).

AGPAT4 partage le plus grand pourcentage d'homologie (75,1%) et d'identité (61,4) avec AGPAT3 (Lu et al., 2005 ; Prasad et al., 2011). A ce jour les seuls travaux publiés portant sur l'analyse de l'activité enzymatique d'Agpat4 sont ceux de Lu et al., 2005. L'activité enzymatique d'Agpat4 a été analysée dans deux modèles différents : un système d'expression *in vitro* utilisant des lysats de réticulocytes de lapin et un système d'expression *in vivo* utilisant la lignée de cellules COS-1. Une activité AGPAT n'a été observée que dans le premier des deux systèmes. Elle était moins de deux fois plus élevée que celle du témoin, alors que l'activité d'Agpat2 était plus de quatre fois plus forte que celle du témoin. L'activité AGPAT d'Agpat4 n'est donc pas bien établie.

### **1.3.2. Profil d'expression des Agpats 1 à 5 et 8**

Le profil d'expression des membres de la famille Agpat a été étudié à l'échelle des ARNm. AGPAT1 est exprimée de façon ubiquitaire (Lu et al., 2005). L'expression d'AGPAT2 a été détectée surtout dans le tissu adipeux viscéral blanc et brun, le foie, le cœur et le pancréas, avec des niveaux modérés dans les muscles, les poumons, l'intestin grêle, les reins, la rate et des niveaux bas dans le cerveau, le placenta, et la graisse sous-cutanée (Eberhardt et al., 1997 ; Lu et al., 2005 ; Vergnes et al., 2006 ; West et al., 1997). AGPAT3 est exprimée de façon ubiquitaire avec un niveau plus élevé dans les testicules, le pancréas et les reins et un niveau légèrement plus bas dans la rate, les poumons, le tissu adipeux et le foie



(Lu et al., 2005 ; Prasad et al., 2011 ; Yuki et al., 2009). Contrairement à AGPAT3, l'expression d'AGPAT4 est spécifique de certains tissus. Son expression est maximale dans le cerveau (Lu et al., 2005 ; Prasad et al., 2011). Le profil d'expression d'AGPAT5 est très similaire à celui d'AGPAT4, mais c'est dans les testicules et la prostate que son expression est la plus forte. AGPAT8 est principalement exprimée dans le cœur, le rein et le muscle squelettique, ainsi que le cerveau, le placenta, et l'intestin grêle (Agarwal et al., 2006 ; Wang et al., 2010). Ces données montrent un recouvrement important des profils d'expression des différentes Agpat.

## **2. Catécholamines**

Nous nous sommes intéressés aux catécholamines, car l'expression d'Agpat4 a été identifiée comme enrichie dans l'aire tegmentaire ventrale et de la substance noire compacte, deux régions catécholaminergiques du cerveau (Brochier et al., 2008). Les catécholamines sont des composés organiques comportant un noyau catéchol (noyau benzénique portant deux groupements hydroxyles adjacents en positions 3 et 4), qui ont des effets physiologiques importants en tant que neurotransmetteurs et hormones. Elles comprennent la dopamine, l'adrénaline et la noradrénaline.

### **2.1. Sites de synthèse et principales fonctions**

#### **2.1.1. La dopamine**

La dopamine, précurseur de l'adrénaline et de la noradrénaline, intervient dans le système nerveux central dans diverses fonctions importantes, telles que le comportement, la cognition, la régulation du mouvement, de l'humeur, de l'attention, et des fonctions végétatives. Sa synthèse a lieu dans les neurones dopaminergiques qui forment dix noyaux dopaminergiques numérotés de A8 à A17 (Marín et al., 2005). Environ 90% de neurones dopaminergiques sont localisés au niveau de trois noyaux mésencéphaliques (A8 - le noyau rétro-rubral, A9 - la substance noire et A10 - l'aire tegmentaire ventrale). Cinq noyaux sont situés dans le diencéphale (A11 à A15) et les autres neurones dopaminergiques sont localisés dans le tubercule olfactif (A16), et dans la rétine (A17) (Marín et al., 2005). Les noyaux dopaminergiques du mésencéphale forment trois systèmes principaux de projections

dopaminergiques: le système nigrostrié, le système mésolimbique et le système mésocortical (Maharajan, 2001).

#### **2.1.1.1. Système nigrostrié**

Le système nigrostrié est constitué des projections dopaminergiques de la substance noire compacte (A9) vers le striatum dorsal (noyau caudé et putamen) et le globus pallidus (Feldman et al., 1997). Cette voie est impliquée dans la boucle motrice décrite dans le chapitre III (Figure 20). Elle participe à la régulation de la motricité volontaire, et sa dégénérescence entraîne la maladie de Parkinson.

La maladie de Parkinson a été décrite par un médecin londonien, Sir James Parkinson, comme « paralysie agitante » (Parkinson, 1817). C'est une maladie fréquente dont les symptômes apparaissent entre 45 et 55 ans pour 2/1000 de la population. Le pourcentage atteint 4/1000 après 85 ans. Les symptômes moteurs se caractérisent par des tremblements plus ou moins prononcés au repos, une lenteur ou absence de mouvements volontaires, une rigidité, une instabilité posturale et des périodes de blocage moteur. Parmi les troubles non moteurs, on retrouve des perturbations olfactives (80% des cas), des troubles gastrointestinaux, vésico-sphinctériens (dans plus de 70% des cas), mais aussi un déficit cognitif. Dans plus de 40% des cas, les patients parkinsoniens souffrent de douleurs chroniques (Ford, 2010).

Les mécanismes impliqués dans le processus neurodégénératif de la maladie de Parkinson ne sont pas encore bien compris. Plusieurs gènes ont été incriminés à partir de l'analyse génétique de rares formes familiales. Parmi ces gènes, certains sont associés à des formes dominantes (alpha-synucléine, LRRk2/dardarine) d'autres à des formes récessives de cette maladie (Parkin, DJ-1) (Harvey et al., 2008 ; Kitada et al., 1998 ; Polymeropoulos et al., 1997). Cependant aucun de ces gènes n'est responsable à lui seul de la maladie.

Les symptômes moteurs de la maladie de Parkinson sont essentiellement la conséquence d'une dégénérescence du système dopaminergique nigro-strié (Bohnen et al., 2007), qui entraîne une dérégulation de la boucle motrice des ganglions de la base. Les neurones dopaminergiques de la voie nigro-striée qui projettent sur les parties dorsolatérales (motrices) du striatum sont plus affectés que ceux qui projettent sur les zones ventrales (cognitives) (Hirsch et al., 1988). Les voies méso-limbique et méso-corticale sont également atteintes. Elles peuvent jouer un rôle dans la formation des troubles cognitifs et certains aspects de l'akinésie (Javoy-Agid and Agid, 1980; Scatton et al., 1982).

Cette maladie est également caractérisée par la présence d'agrégats d'alpha-synucléine formant des inclusions éosinophiles cytoplasmiques intraneuronales, les « corps de Lewy » (Spillantini et al., 1998). Au stade préclinique de la maladie, les inclusions sont majoritairement localisées au niveau du bulbe olfactif et du pont et les patients souffrent de troubles de l'olfaction, de constipation, de troubles des fonctions autonomes et de troubles du sommeil. Ensuite, les corps de Lewy se développent dans la substance noire et on constate l'apparition des troubles moteurs. Dans les derniers stades le néocortex est également atteint ce qui se manifeste par des troubles cognitifs (Braak et al., 2002).

A ce jour il n'existe pas de traitement curatif ni préventif. Le traitement le plus utilisé est la « dopa thérapie » qui consiste à l'administration de L-DOPA (levodopa), le précurseur de la dopamine capable de passer la barrière hémato-encéphalique. Cela permet de compenser la déplétion dopaminergique dans le système nerveux central en rétablissant le taux de la dopamine circulant. En revanche ce traitement n'est pas efficace à long terme, et provoque des complications motrices.

#### **2.1.1.2. Système mésolimbique**

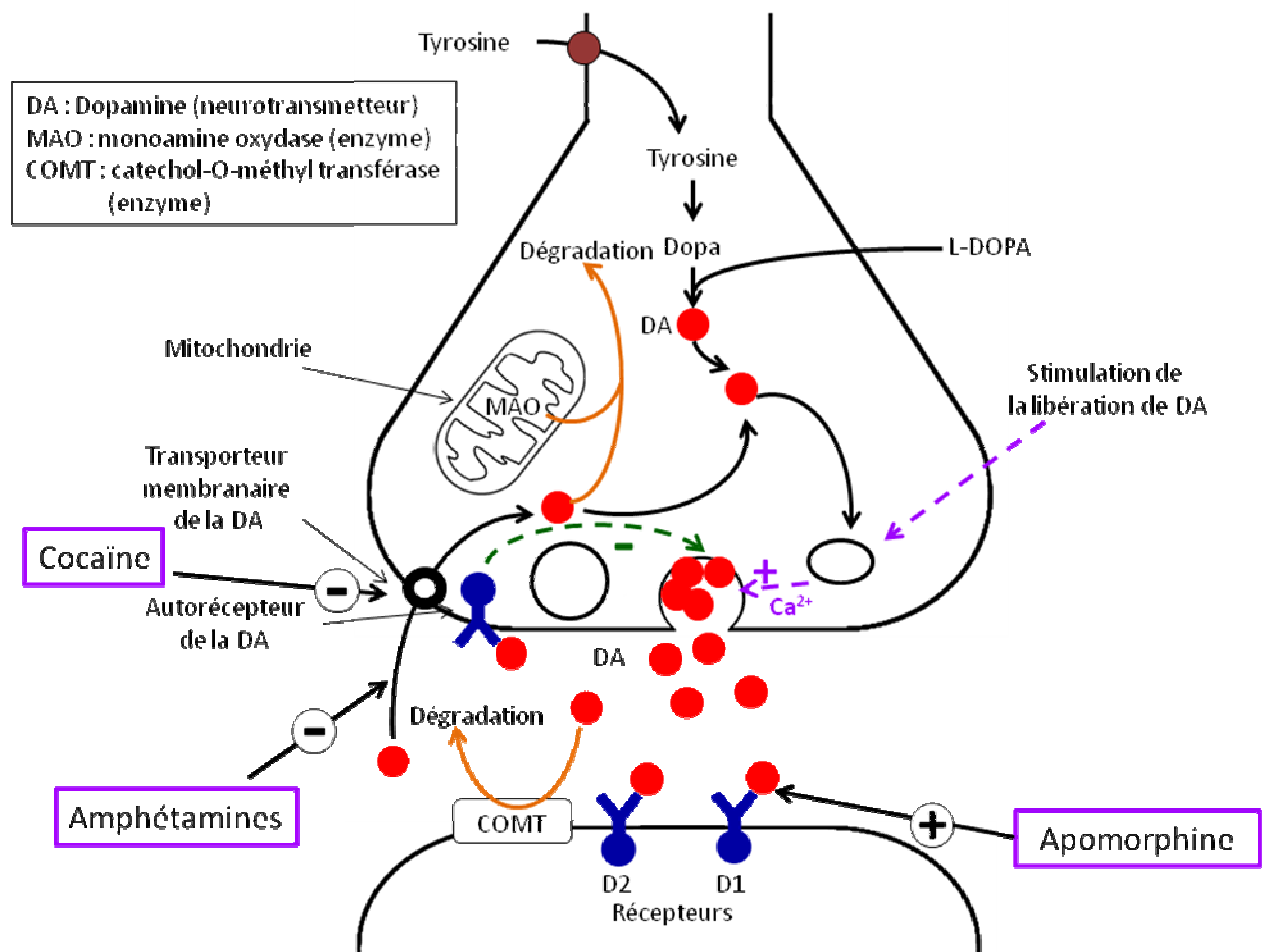
Ce système prend son origine à partir des neurones dopaminergiques de l'aire tegmentaire ventrale (A10) qui projettent sur le noyau accumbens, le bulbe olfactif, ainsi que des structures limbiques telles que le noyau basolatéral de l'amygdale et l'hippocampe (6,7) (Dalhström et Fuxe, 1964). Cette voie participe à de nombreuses fonctions essentielles comme la motivation, la cognition, la mémoire et l'aspect émotionnel de l'apprentissage. Elle est associée à la production des émotions intenses observées chez des patients schizophrènes (van Os et Kapur, 2009), au système de récompense, et à la dépendance aux drogues (Feldman et al., 1997). Le système de récompense est constitué de trois composantes :

- affective : notion de plaisir ou de déplaisir associé à un stimulus
- motivationnelle : motivation pour obtenir les récompenses ou éviter les stimuli aversifs
- cognitive : apprentissage, éducation

La dopamine agit principalement sur la composante motivationnelle et est surtout déterminante dans le repérage de récompenses pour l'organisme. Elle participe en effet à la mémorisation inconsciente d'indices associés aux récompenses. En situation normale, ce système sert à fournir la motivation nécessaire aux comportements de survie (comme l'alimentation et la reproduction), et renforce et conditionne ces comportements par l'intermédiaire de l'activation « de sensations de plaisirs ». Dans l'état addictif, il semble que

le système de récompense soit détourné et modifie le comportement. Ce détournement peut s'expliquer de plusieurs façons suivant le produit (médicaments ou drogues) et son mode de fonctionnement. Ils provoquent tous, de façon directe ou indirecte une augmentation de dopamine dans le cerveau.

Nous pouvons distinguer deux sortes de drogues : les drogues qui agissent directement au niveau de la synapse dopaminergique (cocaïne, amphétamine), et les drogues qui agissent indirectement (opiacés, nicotine, alcool) (Figure 29).



**Figure 29 : La synapse dopaminergique**

Schéma illustrant le devenir de la dopamine dans le neurone présynaptique et dans la synapse. Créée à partir de la tyrosine, la dopamine intègre les vésicules intracytoplasmiques et est excrétée dans la fente synaptique. Elle va alors se fixer sur ses récepteurs (D1 ou D2) puis va être recyclée soit par recapture (recyclage ou élimination par la MAO) soit par élimination par la COMT. Certaines drogues miment les effets de la dopamine (exemple de l'apomorphine) tandis que d'autres empêchent son recyclage et maintiennent l'excitation du neurone post-synaptique plus longtemps (cocaïne bloquant les transporteurs membranaires de la dopamine et amphétamine expulsant la dopamine du neurone présynaptique en faisant fonctionner les transporteurs membranaires de la dopamine en sens inverse).

La consommation de cocaïne est associée à la vigilance, à l'augmentation de l'énergie et de l'activité motrice, à la sensation intense, et à l'euphorie (Preti, 2007). Les effets de renforcement de la cocaïne et ses analogues correspondent généralement à la capacité de la drogue d'inhiber le transporteur de recapture de la dopamine (DAT). Le blocage de cette protéine augmente le niveau synaptique de la dopamine dans le noyau accumbens et autres régions du cerveau essentielles à l'euphorie et à la dépendance (Ritz et al., 1987). La cocaïne augmente aussi les concentrations synaptiques de la sérotonine et de la noradrénaline par le blocage de leurs transporteurs, résultant en effets stimulants et psychotropes. (Bunney et al., 1973 ; Kalivas et Duffy, 1991). Le circuit de dépendance s'installe de façon suivante : le neurone post-synaptique va s'habituer à cette forte concentration en dopamine et synthétiser de nouveaux récepteurs. Cela aura pour effet d'augmenter la sensibilité, ce qui est à l'origine du phénomène de manque (Di Chiara et Imperato, 1988). Le retour à la normale ne se fera qu'après une nouvelle prise de cocaïne de plus en plus importante et fréquente, le neurone s'accoutumant toujours à l'augmentation de la concentration.

Au contraire, l'alcool diminue la sécrétion de dopamine par le maintien de l'excitation des récepteurs GABAergiques, ce qui entraîne l'inhibition des neurones dopaminergiques. Cela a pour effet une somnolence, renforcée par l'action identique de l'alcool sur les récepteurs NMDA du glutamate (Vengeliene et al., 2008). Dans l'alcoolisme chronique, les récepteurs GABAergiques et NMDA deviennent moins sensibles et n'assurent plus leur rôle d'inhibition de la sécrétion de dopamine (Lovinger et al., 1989) ; de surcroît l'alcool diminue l'activité de la monoamine oxydase, qui dégrade la dopamine. Cela se manifeste par une agitation et une irritabilité typiques du manque lié à la dépendance alcoolique.

Quant à la nicotine, elle se fixe sur le récepteur nicotinique de l'acétylcholine, ce qui a pour conséquence une entrée de  $\text{Na}^+$  dans le neurone dopaminergique qui va dépolariser la membrane. Ceci entraînera la libération de dopamine (Pidoplichko et al., 1997). La dépendance s'installe par le fait qu'une fois excité par la nicotine, le récepteur est temporairement désactivé (Pidoplichko et al., 1997). Un fumeur chronique voit alors son plaisir diminué tout au long de la journée. Une légère abstinence comme une nuit de sommeil suffit à réactiver les récepteurs. Il s'ensuit alors une hausse anormale d'activité cholinergique qui se traduit par une agitation et un inconfort qui conduit à une nouvelle cigarette.

### **2.1.1.3. Système mésocortical**

Ce système est composé de projections partant de l'aire tegmentaire ventrale et innervant le cortex préfrontal, le cortex cingulaire antérieur et le cortex entorhinal. Il est impliqué dans la mémorisation, l'attention, et la planification de l'action : le système mésocortical joue ainsi un rôle prépondérant dans les fonctions auxquelles le cortex préfrontal est associé, parmi lesquelles l'activité locomotrice, la régulation des états émotionnels et les processus cognitifs (Goldman-Rakic, 1995 ; Gorenstein et al., 1989 ; Ingvar 1985). Un dysfonctionnement de cette voie pourrait être à l'origine des symptômes de la schizophrénie (comme les hallucinations et les idées délirantes) (van Os et Kapur, 2009). Le blocage pharmacologique de cette voie par des antagonistes dopaminergiques de récepteur D2 (neuroleptiques) réduit les hallucinations mais aussi le fonctionnement global des lobes frontaux.

### **2.1.2. La noradrénaline et l'adrénaline**

La noradrénaline est un neurotransmetteur du système nerveux central ainsi que du système nerveux autonome périphérique sympathique. Elle est synthétisée dans les neurones noradrénergiques qui forment au niveau du tronc cérébral sept noyaux: A1 à A3 dans le bulbe rachidien ; A4 à A7 dans le pont ; A6, le locus coeruleus, est le noyau noradrénergique le plus important (Wallace et al., 1992). Les axones issus du locus coeruleus innervent presque toutes les parties du cerveau : le cortex cérébral, le thalamus, l'hypothalamus, les bulbes olfactifs, le cervelet, le mésencéphale, et la moelle épinière. Le locus coeruleus est impliqué dans les processus attentionnels, l'éveil, et les cycles veille-sommeil, le contrôle cardiovasculaire, ainsi que dans l'apprentissage et la mémoire, l'anxiété, la douleur, l'humeur et le métabolisme du cerveau (Sasaki et al., 2008).

Le système sympathique participe à la mise en état d'alerte de l'organisme et à la préparation à l'activité physique et intellectuelle. La noradrénaline est synthétisée également au niveau de la médullosurrénale mais en plus faible quantité. Son action s'exerce en se fixant sur les récepteurs alpha et bêta adrénergiques, et s'effectue entre autres en : facilitant la transmission de l'influx nerveux ; entraînant la vasoconstriction artérielle ; augmentant la pression artérielle (Choudhury et al., 2009).

La noradrénaline est un précurseur de l'adrénaline (connue aussi sous le nom d'épinéphrine). L'adrénaline est synthétisée au niveau du tronc cérébral par les groupes de

cellules noradrénergiques/adrénergiques (A1/C1, A2/C2) et adrénergiques (C3). Le noyau A1/C1 est le noyau le plus caudal du tronc cérébral et représente environ 70% de neurones adrénergiques du système nerveux central (Howe et al., 1980). Le noyau A2/C2 est situé au niveau du noyau solitaire et représente environ 10% des neurones adrénergiques du système nerveux central (Howe et al., 1980). Le noyau C3 représente environ 10% des neurones adrénergiques du système nerveux central.

L'adrénaline est aussi synthétisée dans la médullosurrénale par les cellules chromaffines. Celles-ci libèrent l'adrénaline dans le sang ce qui permet d'assurer une réponse viscérale coordonnée. Elle entraîne une accélération du rythme cardiaque, de la pression artérielle, dilate les bronches et les pupilles et élève la concentration du glucose sanguin (Schneider et al., 2011).

## **2.2. Métabolisme des catécholamines**

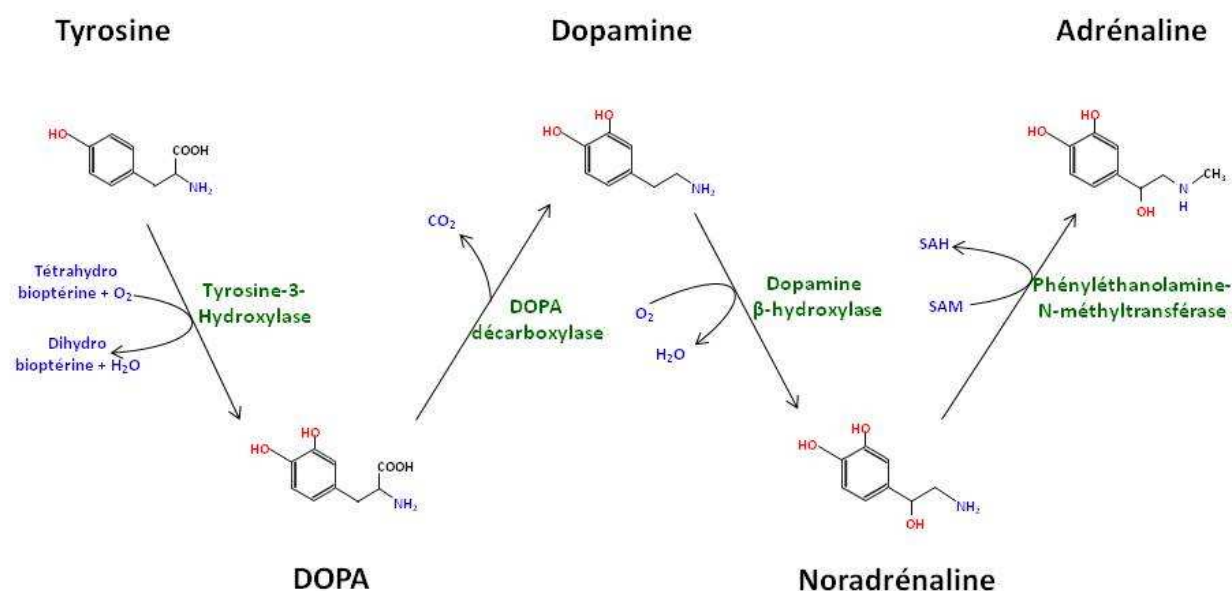
### **2.2.1. Biosynthèse**

Tous les neurones catécholaminergiques contiennent la tyrosine hydroxylase (TH), l'enzyme catalysant la première réaction de la biosynthèse des catécholamines : la transformation de la tyrosine en L-DOPA (L-dihydroxyphénylalanine). L'activité de la TH est limitante pour la biosynthèse des catécholamines (Nagatsu et al., 1989). Son activité est régulée par des signaux provenant du cytoplasme de la terminaison axonique. L'augmentation de la concentration de dopamine entraîne une inhibition de l'activité de la TH. L'augmentation de la concentration des ions calcium au niveau de la synapse stimule l'activité de la TH.

La DOPA est ensuite convertie en dopamine par la DOPA décarboxylase, qui est présente en quantité élevée dans les neurones dopaminergiques de sorte que la quantité de la dopamine synthétisée ne dépend que de la disponibilité de la DOPA. Une fois synthétisée, la dopamine s'accumule dans des vésicules synaptiques et va être libérée au niveau des terminaisons axonales par exocytose suite à une entrée de calcium provoquée par un potentiel d'action. Cette libération synaptique permet d'activer des récepteurs dopaminergiques post-synaptiques, mais aussi des récepteurs situés sur les terminaisons dopaminergiques, dits autorécepteurs, qui ont une action régulatrice sur la libération de dopamine.

Pour synthétiser la noradrénaline, la dopamine est prise en charge par l'enzyme dopamine  $\beta$ -hydroxylase. La synthèse de la noradrénaline a lieu dans les vésicules

synaptiques. La noradrénaline est le précurseur de l'adrénaline, qui est synthétisée par la phényléthanolamine N-méthyltransférase présente au niveau du cytosol des cellules adrénergiques (Figure 30)



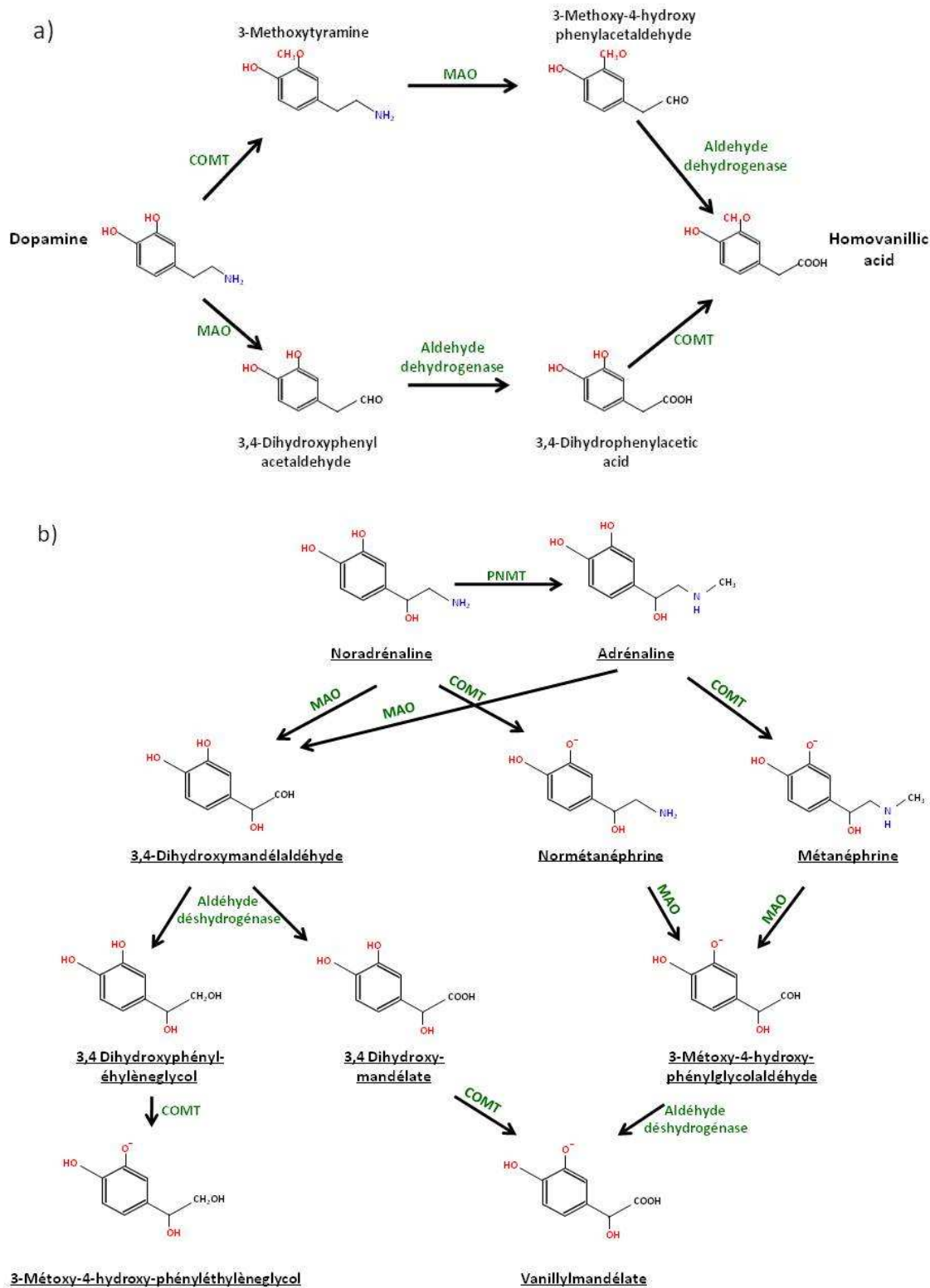
**Figure 30 : Biosynthèse des catécholamines**  
Illustrations des différentes étapes de la synthèse des catécholamines

### 2.2.2. Dégradation

La dopamine est rapidement recapturée par le transporteur (DAT) situé sur la membrane des terminaisons présynaptiques (Jaber et al., 1997). Les DAT sont exprimés sur la majorité des neurones dopaminergiques, et leur nombre influence la durée d'action de la dopamine sur ses récepteurs (Ciliax et al., 1999). Après recapture elle peut soit être réincorporée dans des vésicules synaptiques et être réutilisée, ou bien elle est dégradée par la monoamine oxydase (MAO) qui la convertit en 3,4-dihydroxyphénylaldéhyde (DOPAL) puis l'aldéhyde déshydrogénase transforme la DOPAL en acide 3,4-dihydroxyphenylacétique (DOPAC). Cette dernière est ensuite dégradée en acide homovanillique (HVA) dans le milieu extracellulaire grâce à l'action de la catéchol-O-méthyltransférase (COMT). La COMT peut aussi dégrader directement la dopamine extracellulaire en 3-méthoxytyramine (3-MT) (Kirov et al., 1998; Benoit-Marand et al., 2000) (Figure 31).

L'adrénaline et la noradrénaline sont également dégradées par l'action de la MAO et de la COMT (Figure 31).





**Figure 31 : Dégradation des catécholamines (D'après KEGG)**

a) Deux voies de dégradation de la dopamine initiées soit par l'enzyme mitochondriale MAO soit par l'ectoenzyme COMT. La dégradation par MAO se réalise dans le neurone présynaptique. La COMT agit dans la fente synaptique

b) Dégradation de la noradrénaline et de l'adrénaline utilisant les mêmes enzymes (MAO et COMT). Deux résidus communs aux deux catécholamines sont créés : la Vanillylmandélate et la 3-Méthyl-4-hydroxyphényléthylèneglycol.

### 3. Article N°2

Avant mon arrivée en thèse une lignée de souris déficientes pour le gène d'Agpat4 a été construite au laboratoire à partir de cellules ES provenant de l'institut Sanger. Chez ces souris, le gène d'Agpat4 est invalidé par l'insertion d'une cassette de « *gene trap* » dans le premier intron, en amont du codon ATG situé dans le deuxième exon. Au cours de ma thèse, j'ai localisé précisément le site d'insertion de la cassette de « *gene trap* » par séquençage de l'ADN génomique. Cela m'a permis ensuite de mettre en place une stratégie de génotypage des souris transgéniques par PCR multiplex. J'ai également vérifié par Southern blot qu'une seule copie de la cassette s'est intégrée dans le génome de ces souris et, par Northern blot, que l'expression d'Agpat4 est efficacement invalidée. J'ai analysé le profil d'expression du gène rapporteur codant la  $\beta$ -galactosidase, présent dans la cassette de « *gene trap* », chez les souris transgéniques, ainsi que le profil d'expression d'Agpat4 chez les souris sauvages par RT-PCR quantitative.

Une éventuelle compensation de l'absence d'Agpat4 chez les souris transgéniques par d'autres membres de la famille Agpat a été envisagée. J'ai ainsi procédé à la quantification des ARNm des paralogues suivants : Agpat3, Agpat5 et Agpat8 (les paralogues les plus proches d'Agpat4) ; Agpat1 et Agpat2 (les Agpats les mieux caractérisées) ; Agpat7 (une lysophosphatidyléthanolamine acyltransférase) par qRT-PCR chez les souris transgéniques comparées aux souris sauvages. Agpat4 étant une acyltransférase putative, nous avons voulu évaluer les conséquences métaboliques de son invalidation. Une approche métabolomique *in vivo* a été employée. Les métabolomes de tissus dans lesquels Agpat4 est exprimée ont été comparés entre les souris sauvages et transgéniques. Une analyse non ciblée par spectrométrie de masse couplée à la chromatographie en phase liquide nous a permis de détecter des métabolites provenant de différentes voies métaboliques perturbés par l'absence d'Agpat4. Plusieurs de ces molécules ont pu être identifiées grâce à la technique MS/MS.

## **Agpat4 influences the metabolism of lysophosphatidylethanolamines and catecholamines *in vivo***

Nad'a Lepejová<sup>1</sup>, Camille Brochier<sup>1</sup>, Geoffrey Madalinski<sup>2</sup>, Laurie Galvan<sup>3</sup>, Jérôme Cotton<sup>2</sup>, Marie-Claude Gaillard<sup>3</sup>, Ying Xu<sup>4</sup>, Patrick Héry<sup>1</sup>, Nicole Faucon-Biguët<sup>5</sup>, Jean-François Heilier<sup>4,6</sup>, Matthieu Gérard<sup>1</sup>, Jean-Marc Elalouf<sup>1</sup>, Christophe Junot<sup>4</sup>, Michel de Chaldée<sup>1,\*</sup>

<sup>1</sup> CEA, iBiTecS, Service de Biologie Intégrative et Génétique Moléculaire (SBIGeM), 91191 Gif-sur-Yvette Cedex, France; <sup>2</sup> Profilomic, CEA Saclay, 91191 Gif-sur-Yvette Cedex, France; <sup>3</sup> CEA, I2BM, Molecular Imaging Research Center (MIRCen), URA CEA-CNRS 2210, 92265 Fontenay-aux-Roses, France; <sup>4</sup> CEA, iBiTecS, Service de Pharmacologie et d'Immunoanalyse (SPI), 91191 Gif-sur-Yvette Cedex, France; <sup>5</sup> CNRS UMR 7091, Hôpital de la Pitié-Salpêtrière, 75013 Paris, France; <sup>6</sup> Université catholique de Louvain, Louvain Centre for Toxicology and Applied Pharmacology (LTAP), 1200 Brussels, Belgium.

\* To whom correspondence may be addressed. E-mail: [michel.de.chaldee@cea.fr](mailto:michel.de.chaldee@cea.fr); tel: +33 1 6908 8865; fax: +33 1 6908 4712.

The authors declare no conflict of interest.

## Abstract

Agpat4 (1-acylglycerol-3-phosphate *O*-acyltransferase 4) has been classified into a family of *O*-acyltransferases involved in the biosynthesis of glycerophospholipids from sequence homologies, but its physiological role is unknown. Its highest expression levels are found in the brain, in particular in dopaminergic structures of the midbrain. Here, we further show that it is expressed in most central and peripheral catecholaminergic tissues. To investigate the function of Agpat4 *in vivo*, we generated Agpat4-deficient mice and compared the metabolic profiles of their ventral midbrain, striatum and adrenal gland to those of wild-type littermates using ultra-high performance liquid chromatography coupled with mass spectrometry. Several classes of metabolites were perturbed by Agpat4 deficiency. Lysophosphatidylethanolamines were broadly affected, revealing that Agpat4 may act as a lysophosphatidylethanolamine acyltransferase *in vivo*. Monogalactosylalkylacylglycerol was the most upregulated metabolite in Agpat4-deficient midbrains, indicating, together with the downregulation of a possible galactocerebroside, that Agpat4 may be involved in the control of myelin lipid composition. Non-lipid metabolites were also affected. Notably, changes in catecholamine metabolism, in particular in dopamine and norepinephrine levels, were observed in the striatum and adrenal gland of Agpat4-deficient mice, suggesting a possible role of Agpat4 in catecholamine physiology and connecting Agpat4 function to its expression pattern.

**Keywords:** Agpat; Substantia nigra; Basal ganglia; Metabolome; Phosphatidylethanolamine; Galactosyl glycerol; Gene trap

## Introduction

Agpat4 (1-acylglycerol-3-phosphate *O*-acyltransferase 4) has been named after sequence homologies with members of a family of *O*-acyltransferases involved in the biosynthesis of glycerophospholipids, a class of lipids important not only for membrane structure but also as precursors of lipid mediators (Frisardi et al., 2011). This family comprises glycerol-3-phosphate *O*-acyltransferases (Gpats) and *bona fide* Agpats, whose activities are essential to the *de novo* biosynthesis of all glycerophospholipids, and other lysophospholipid acyltransferases, which are involved in the turnover of the fatty acid moiety at the *sn*-2 position of glycerophospholipids (Hishikawa et al., 2008; Shindou and Shimizu, 2009). Several enzymes termed Agpats from sequence homologies have subsequently been reported to exhibit other activities. For instance, Agpat6 (Nagle et al., 2008) and Agpat9 (Cao et al., 2006) turned out to be Gpats. The Agpat activity of Agpat4 has been tested in a single study, both in reticulocyte lysates, where its activity was less than twice that of the control and less than half that of Agpat2, and in transiently transfected cells overexpressing Agpat4, where no Agpat activity could be detected (Lu et al., 2005a). Thus the reactions actually catalyzed by Agpat4 remain unclear.

Scattered information on Agpat4, in particular on its mRNA tissue distribution, is included in publications on other Agpats (Lu et al., 2005a; Lu et al., 2005b; Beigneux et al., 2006; Vergnes et al., 2006; Prasad et al., 2011), but so far it has not been the subject of a specific study. These data reveal that Agpat4 has a distinctive expression pattern among Agpats, with highest mRNA levels in the brain and low levels in both white and brown adipose tissues. Within the brain, we have observed that Agpat4 expression is enriched in midbrain dopaminergic nuclei (Brochier et al., 2008), notably in the substantia nigra pars compacta, which is selectively affected in Parkinson's disease.

To investigate the physiological function of Agpat4 and the significance of its unique tissue distribution, we generated an Agpat4-deficient mouse line and compared the metabolic profiles of brain regions (the ventral midbrain and the striatum) and adrenal tissue between Agpat4-deficient and wild-type littermates using ultra-high performance liquid chromatography coupled with mass spectrometry. Comparable untargeted metabolomic approaches have been successfully employed to identify endogenous substrates and possible physiological functions for lipid metabolism enzymes in the whole mouse brain (Saghatelian et al., 2004; Kopp et al., 2010). Here, we show that, despite no overt behavioral or physiological abnormalities, Agpat4-deficient mice exhibit a metabolic phenotype. Different classes of metabolites were affected by Agpat4 deficiency, notably lysophospholipids, lipid components of myelin and catecholamines. This work provides information on the biochemical activity of Agpat4 *in vivo* and relates its activity to the functions of the tissues where it is expressed.

## Results

**Generation of Agpat4-deficient mice.** To analyze the function of Agpat4 *in vivo*, we derived Agpat4-deficient mice from a gene-trap ES cell line. Homozygous mice were obtained in the expected Mendelian ratio. They are apparently healthy and their behavior is normal at the gross level. Both males and females are fertile. The gene-trap cassette integrated into the first intron of the *Agpat4* gene, upstream of the translation initiation codon (Fig. 1 A and B), causing a dramatic reduction in the amount of Agpat4-encoding transcripts in these mice (Fig. 1C).

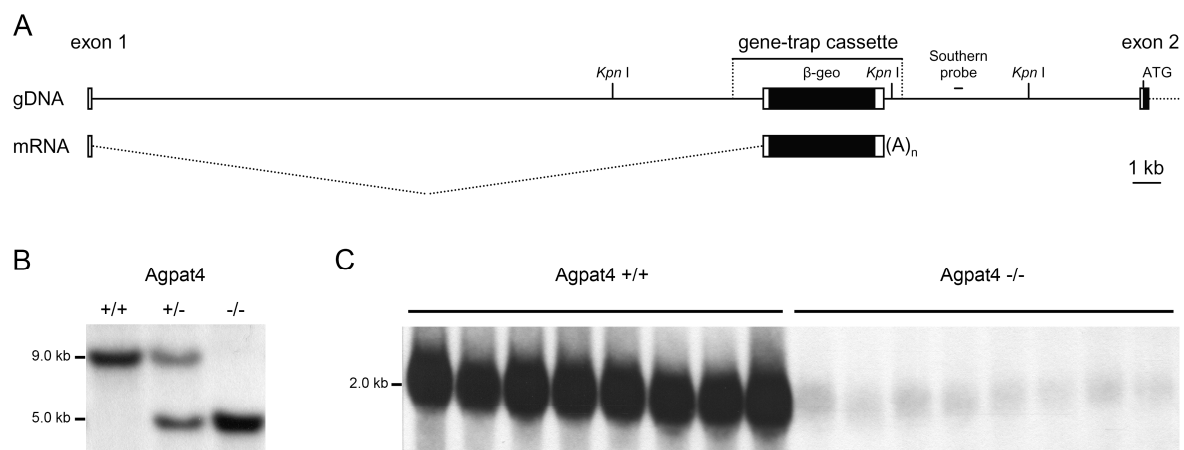


Fig. 1

**Fig. 1.** Molecular characterization of *Agpat4*-deficient mice. (A) Trapped *Agpat4* allele and corresponding transcript. Transcribed sequences are depicted by boxes and coding sequences are shown in black. (B) Southern hybridization profile of *Kpn* I-digested genomic DNA from wild-type, heterozygous and *Agpat4*-deficient littermates using the probe shown in (A). (C) Northern blot analysis of 25- $\mu$ g total RNA samples from the brains of eight wild-type and eight *Agpat4*-deficient littermates using a probe encompassing the entire *Agpat4* coding sequence. A single signal was detected in each lane at approximately 2 kb, consistent with the expected size of full-length *Agpat4* transcripts. Detailed probe information is provided in Table S1.  $\beta$ -geo, beta-galactosidase-neomycin phosphotransferase fusion gene.

To assess whether the lack of a conspicuous phenotype could be attributable to functional compensation by a related protein, we measured the mRNA levels of Agpat family members (Agpat3, Lclat1 and Agpat5, the closest paralogs of Agpat4; Agpat1 and Agpat2, two well-characterized lysophosphatidic acid acyltransferases; Lpcat4, a lysophosphatidylethanolamine acyltransferase) in Agpat4-deficient mice and wild-type littermates. None of these transcripts was significantly upregulated in Agpat4-deficient samples (Fig. S1).

**Agpat4 is expressed in catecholaminergic tissues.** Agpat4 expression can be detected in several brain regions. Using quantitative methods, we have previously shown that its transcripts are particularly enriched in the substantia nigra and ventral tegmental area, two midbrain dopaminergic structures (Brochier et al., 2008). To assess whether Agpat4 expression is a common feature of catecholaminergic tissues, we took advantage of the beta-galactosidase reporter gene included in the gene-trap cassette. As expected, X-Gal staining was intense in the substantia nigra pars compacta and ventral tegmental area of Agpat4-deficient mice (Fig. 2A). Strikingly, most central (Fig. S2) and peripheral (Fig. 2B) sites of catecholamine synthesis were stained.



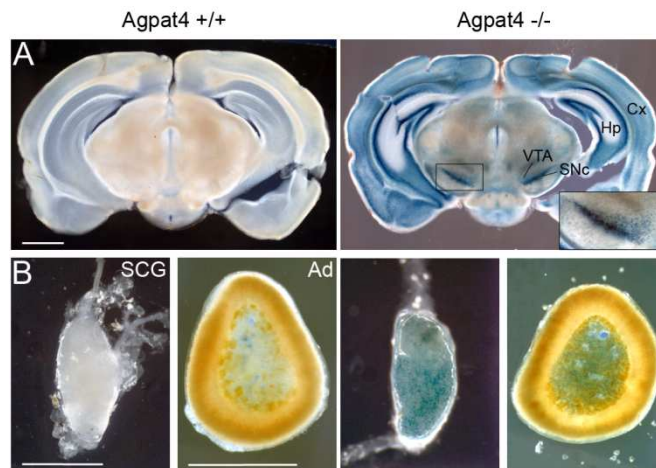


Fig. 2

**Fig. 2.** Beta-galactosidase reporter gene expression in *Agpat4*-deficient mice. (A) X-Gal staining of 300- $\mu$ m coronal brain sections from wild-type and *Agpat4*-deficient littermates at the level of the substantia nigra pars compacta (SNc) and ventral tegmental area (VTA). The cortex (Cx) and hippocampus (Hp) are also stained. (B) Staining of whole superior cervical ganglia (SCG; noradrenergic) and 300- $\mu$ m adrenal gland (Ad) sections showing the medulla (noradrenergic and adrenergic). (Scale bars, 1 mm.)

**Lysophosphatidylethanolamine levels are altered in *Agpat4*-deficient tissues.** To investigate the function of *Agpat4* *in vivo*, we compared the metabolic profiles of the ventral midbrain, the striatum, which is the target of nigral dopaminergic neurons, and the adrenal gland between *Agpat4*-deficient ( $n = 8$ ) and wild-type littermates ( $n = 8$ ). This analysis was performed twice independently on different samples. For each tissue, we selected the metabolite signals that were consistently found to be among the 100 most up- or downregulated in *Agpat4*-deficient mice in both profiling experiments.

Using this criterion, we observed that glycerophospholipid metabolism was altered in all *Agpat4*-deficient tissues (Table 1). Lysophosphatidylethanolamines (LPEs) were especially affected: LPEs with short ( $\leq 17$  carbons) fatty acid chains were mostly upregulated in *Agpat4*-deficient mice while LPEs with medium (18-20 carbons) fatty acid chains were downregulated (Fig. 3). No major modifications in diacylphosphatidylethanolamine levels

could be detected (Fig. S4). Since Agpat4 has been described as a probable lysophosphatidic acid acyltransferase (Lu et al., 2005a), we searched our data for putative phosphatidic acids, but few were found (Fig. S4).

**Table 1. Glycerophospholipids affected in Agpat4-deficient mice**

| Upregulated metabolites       |   | m/z; retention time  | Evidence       | Tissue                      |
|-------------------------------|---|--|----------------|-----------------------------|
| Lysophosphatidylethanolamines |   |  |                |                             |
| LPE(14:0)                     | C <sub>19</sub> H <sub>40</sub> NO <sub>7</sub> P | 424.2439 [M-H] <sup>-</sup> ; 10.1 min   | MS/MS          | midbrain, striatum          |
| LPE(16:1)                     | C <sub>21</sub> H <sub>42</sub> NO <sub>7</sub> P | 450.2597 [M-H] <sup>-</sup> ; 10.4 min<br>452.2772 [M+H] <sup>+</sup> ; 10.1 min | MS/MS<br>MS/MS | midbrain, striatum          |
| Downregulated metabolites     |   | m/z ; retention time   | Evidence       | Tissue                      |
| Lysophosphatidylethanolamines |   |  |                |                             |
| LPE(18:2)                     | C <sub>23</sub> H <sub>44</sub> NO <sub>7</sub> P | 476.2757 [M-H] <sup>-</sup> ; 10.8 min   | Kendrick       | midbrain                    |
| LPE(20:1)                     | C <sub>25</sub> H <sub>50</sub> NO <sub>7</sub> P | 506.3227 [M-H] <sup>-</sup> ; 12.3 min<br>508.3410 [M+H] <sup>+</sup> ; 11.9 min | MS/MS          | midbrain, striatum          |
| LPE(20:2)                     | C <sub>25</sub> H <sub>48</sub> NO <sub>7</sub> P | 504.3067 [M-H] <sup>-</sup> ; 11.7 min   | MS/MS          | midbrain, striatum          |
| LPE(20:3)                     | C <sub>25</sub> H <sub>46</sub> NO <sub>7</sub> P | 502.2914 [M-H] <sup>-</sup> ; 11.2 min<br>504.3075 [M+H] <sup>+</sup> ; 11.0 min | MS/MS          | midbrain, striatum, adrenal |
| LPE(22:4)                     | C <sub>27</sub> H <sub>48</sub> NO <sub>7</sub> P | 528.3072 [M-H] <sup>-</sup> ; 11.6 min<br>530.3251 [M+H] <sup>+</sup> ; 11.2 min | MS/MS          | adrenal                     |
| LPE(22:4-OH)                  | C <sub>27</sub> H <sub>48</sub> NO <sub>8</sub> P | 544.3001 [M-H] <sup>-</sup> ; 10.9 min<br>546.3188 [M+H] <sup>+</sup> ; 10.5 min | MS/MS          | midbrain                    |
| Other glycerophospholipids    |   |  |                |                             |
| LPI(18:1)                     | C <sub>27</sub> H <sub>51</sub> O <sub>12</sub> P | 597.3012 [M-H] <sup>-</sup> ; 11.2 min   | MS/MS          | midbrain                    |
| LPS(22:4)                     | C <sub>28</sub> H <sub>48</sub> NO <sub>9</sub> P | 572.2959 [M-H] <sup>-</sup> ; 11.5 min   | putative       | midbrain, striatum          |
| LPG(22:6)                     | C <sub>28</sub> H <sub>45</sub> O <sub>9</sub> P  | 555.2689 [M-H] <sup>-</sup> ; 11.1 min   | MS/MS          | adrenal                     |

Up- or downregulation was assessed by calculating the ratio of the mean signal intensity in Agpat4-deficient mice ( $n = 8$ ) to that in wild-type littermates ( $n = 8$ ). Metabolite identification was inferred from MS/MS (Fig. S3) or Kendrick mass defect analysis (Table S2; Hughey et al., 2001).

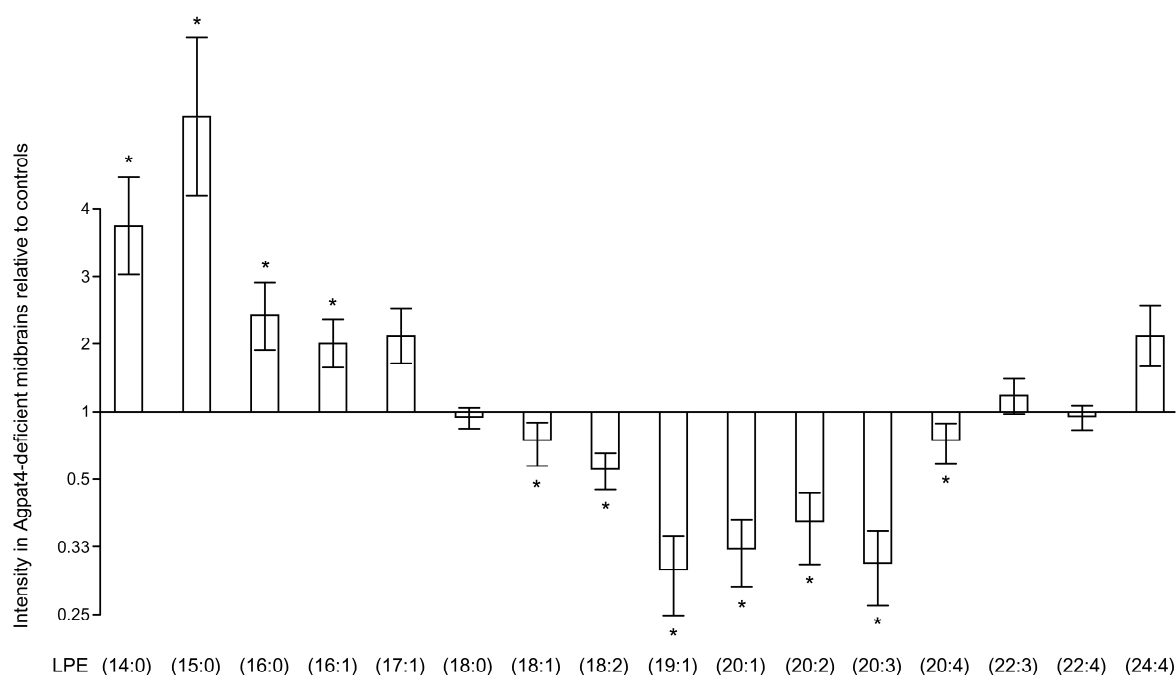


Fig. 3

**Fig. 3.** Lysophosphatidylethanolamine levels in the ventral midbrain of Agpat4-deficient mice. Signals corresponding to lysophosphatidylethanolamines (LPEs) were extracted from profiling data by Kendrick mass defect analysis (Table S2; Hughey et al., 2001). Bars represent the average ratio  $\pm$  SEM of the signal intensity in Agpat4-deficient mice ( $n = 8$ ) to the mean signal intensity in wild-type littermates ( $n = 8$ ). Results are shown for one profiling experiment for  $[M-H]^-$  signals whose intensity was greater than  $10^5$  in at least three samples of either genotype. The data were mean-centered and scaled to Pareto variance before OPLS.

\* VIP score  $> 3$ .

**Agpat4 deficiency perturbs the metabolism of several lipid classes and of catecholamines.** Agpat4 deficiency had an effect not only on glycerophospholipids but also on other lipid classes (Table 2). As for glycerophospholipids, the species upregulated in Agpat4-deficient mice had short fatty acid chains, while the downregulated species had longer fatty acid chains. Notably, monogalactosylalkylacylglycerol (0-16:0/16:0) was the most upregulated metabolite in both profiling experiments, with intensities 8.3 and 5.4 times higher in Agpat4-deficient than in wild-type midbrains. Inversely, several sphingolipids, including monoglycosylceramide (36:1), a possible galactocerebroside, were found to be downregulated in Agpat4-deficient midbrains.

**Table 2. Other metabolites affected in Agpat4-deficient mice**

| Upregulated metabolites        |   | m/z; retention time                       | Evidence | Tissue   |
|--------------------------------|---|---|----------|----------|
| Monogalactosylglycerols        |   |   |          |          |
| monoacyl (0:0/16:0)            | C <sub>26</sub> H <sub>50</sub> O <sub>11</sub>             | 537.3253 [M+FA-H] <sup>-</sup> ; 11.8 min | Kendrick | midbrain |
| alkylacyl (O-16:0/16:0)        | C <sub>42</sub> H <sub>82</sub> O <sub>11</sub>             | 761.5738 [M+FA-H] <sup>-</sup> ; 20.1 min | MS/MS    | midbrain |
| Other metabolites              |   |   |          |          |
| Adenine                        | C <sub>5</sub> H <sub>5</sub> N <sub>5</sub>                | 136.0617 [M+H] <sup>+</sup> ; 1.1 min     | standard | adrenal  |
| Norepinephrine                 | C <sub>8</sub> H <sub>11</sub> NO <sub>3</sub>              | 168.0661 [M-H] <sup>-</sup> ; 1.1 min     | MS/MS    | adrenal  |
| Dihydroxyphenylglycol          | C <sub>8</sub> H <sub>10</sub> O <sub>4</sub>               | 169.0498 [M-H] <sup>-</sup> ; 1.5 min     | putative | adrenal  |
| Downregulated metabolites      |   | m/z; retention time                       | Evidence | Tissue   |
| Sphingolipids                  |   |   |          |          |
| Sphingosine-1-phosphate(d18:1) | C <sub>18</sub> H <sub>38</sub> NO <sub>5</sub> P           | 378.2395 [M-H] <sup>-</sup> ; 10.3 min    | putative | midbrain |
| Sphinganine-1-phosphate(d18:0) | C <sub>18</sub> H <sub>40</sub> NO <sub>5</sub> P           | 380.2553 [M-H] <sup>-</sup> ; 10.5 min    | putative | midbrain |
| Psychosine sulfate(d18:1)      | C <sub>24</sub> H <sub>47</sub> NO <sub>10</sub> S          | 540.2815 [M-H] <sup>-</sup> ; 10.4 min    | putative | midbrain |
| Monoglycosylceramide(36:1)     | C <sub>43</sub> H <sub>83</sub> NO <sub>10</sub>            | 772.5917 [M+FA-H] <sup>-</sup> ; 17.8 min | MS/MS    | midbrain |
| Other metabolites              |   |   |          |          |
| Dihydrothymine                 | C <sub>5</sub> H <sub>8</sub> N <sub>2</sub> O <sub>2</sub> | 127.5658 [M+H] <sup>+</sup> ; 1.0 min     | putative | striatum |
| Acetylcholine                  | C <sub>7</sub> H <sub>16</sub> NO <sub>2</sub>              | 146.1173 [M+H] <sup>+</sup> ; 1.0 min     | MS/MS    | striatum |
| Dopamine                       | C <sub>8</sub> H <sub>11</sub> NO <sub>2</sub>              | 154.0858 [M+H] <sup>+</sup> ; 1.0 min     | standard | striatum |
| Phospho-γ-gluconolactone       | C <sub>6</sub> H <sub>11</sub> O <sub>9</sub> P             | 257.0058 [M-H] <sup>-</sup> ; 1.2 min     | MS/MS    | adrenal  |
| unassigned                     | -   | 495.3058 [M-H] <sup>-</sup> ; 14.0 min    | -        | striatum |

Comments as for Table 1. Adenine and dopamine were identified by comparison with authentic standards. FA, formic acid.

Apart from lipids, a number of small molecules were also affected by Agpat4 deficiency (Table 2). In particular, several catecholamine-related metabolites showed altered levels in Agpat4-deficient tissues (Fig. 4), including the neurotransmitters dopamine (detected

in the striatum but synthesized by midbrain neurons) and norepinephrine (detected in the adrenal gland).

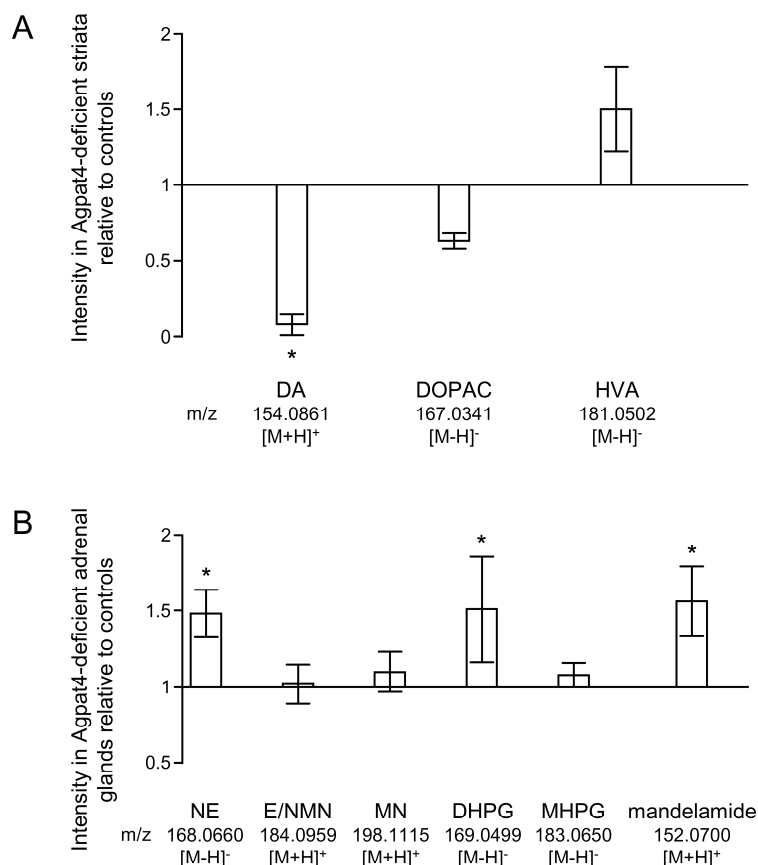


Fig. 4

**Fig. 4.** Levels of catecholamine-related metabolites in Agpat4-deficient tissues. Bars represent the average ratio  $\pm$  SEM of the signal intensity in Agpat4-deficient mice ( $n = 8$ ) to the mean signal intensity in wild-type littermates ( $n = 8$ ). Results are shown for one profiling experiment. (A) Dopamine metabolites detected in the striatum. Dopamine (DA) was identified by comparison with an authentic standard. 3,4-Dihydroxyphenylacetic acid (DOPAC) and homovanillic acid (HVA) are putative. The data were mean-centered and scaled to Pareto variance before PLS-DA. \* VIP score  $> 3$ . (B) Norepinephrine and epinephrine metabolites detected in the adrenal gland. Epinephrine or normetanephrine (E/NMN) and metanephrine (MN) were identified by comparison with authentic standards, and norepinephrine (NE) and mandelamide from MS/MS analysis. Dihydroxyphenylglycol (DHPG) and 3-methoxy-4-hydroxyphenylglycol (MHPG) are putative. The data were mean-centered and scaled to unit variance before PLS-DA. \* VIP score  $> 1$ .

## Discussion

In this study, we generated mice deficient for Agpat4, a putative lysophospholipid acyltransferase mainly expressed in the brain, and analyzed their metabolic phenotype.

We observed that glycerophospholipid levels were altered in all Agpat4-deficient tissues examined. Although Agpat4 has been postulated to be a lysophosphatidic acid acyltransferase (Lu et al., 2005a), we found that most of the affected glycerophospholipids were lysophosphatidylethanolamines (LPEs). A systematic screening of LPEs in Agpat4-deficient mice indicated that LPEs with short fatty acid chains were upregulated and LPEs with medium chains were downregulated. Since short saturated chains tend to be more common in the *sn*-1 position and long unsaturated chains in the *sn*-2 position (Lands, 2000), our results suggest that LPEs acylated in the *sn*-1 position are upregulated in Agpat4-deficient mice whereas those acylated in the *sn*-2 position are downregulated. LPEs acylated in the *sn*-2 position are likely to originate from the hydrolysis of diacylphosphatidylethanolamines by phospholipase A1 (Pete et al., 1994). Therefore, we propose that, although the few putative diacylphosphatidylethanolamines that could be detected in our dataset were not downregulated in Agpat4-deficient tissues, Agpat4 may catalyze the acylation of the *sn*-2 position of *sn*-1-acyl LPEs. Other members of the Agpat family have been reported to have a lysophosphatidylethanolamine acyltransferase activity, notably Lpcat4 (also termed Agpat7 or Lpeat2; Cao et al., 2008) and Agpat5 (Prasad et al., 2011). Both are expressed in the brain but, from *in situ* hybridization (Lein et al., 2007) and transcriptome data (Brochier et al., 2008) in adult mice, their transcripts do not have the same regional distribution within the brain as Agpat4 transcripts. Here, we showed that Agpat4 deficiency was not compensated for by Lpcat4 or Agpat5 overexpression in the ventral midbrain, but the lysophosphatidylethanolamine acyltransferase activity is highly redundant (Hishikawa et al., 2008).

The effect of Agpat4 deficiency on other glycerophospholipids may be accounted for by connections between their biosynthetic pathways (Shindou and Shimizu, 2009). It is also conceivable that Agpat4 could catalyze the acylation of a variety of glycerophospholipids, as described for other members of the Agpat family (Yuki et al., 2009; Prasad et al., 2011; Zhao et al., 2009). In particular, very few putative phosphatidic acids were detected in our dataset, so we cannot exclude that Agpat4 may have a lysophosphatidic acid acyltransferase activity.

Other lipid classes were found to be affected by Agpat4 deficiency. Strong upregulation of monogalactosylalkylacylglycerol (0-16:0/16:0) was consistently observed in the ventral midbrain of Agpat4-deficient mice. It has been characterized in the mammalian brain (Pahlsson et al., 2001), where it is considered, like other galactosylglycerols, as a minor component of myelin (Pieringer et al., 1977). It is abundant in the testis, where it is the precursor (Honke et al., 1996) and degradation product (Gadella et al., 1993) of seminolipid, the major glycolipid of mammalian spermatozoa. The final step of monogalactosylalkylacylglycerol synthesis is catalyzed by ceramide galactosyltransferase (Fujimoto et al., 2000), which is also responsible for the biosynthesis of galactocerebrosides, the most abundant lipids in myelin. Interestingly, a monoglycosylceramide (36:1), possibly a galactocerebroside, was also affected by Agpat4 deficiency. These observations suggest that Agpat4 may influence myelin lipid composition and properties. This effect could result from connections between glycerophospholipid, galactosylglycerol and sphingolipid metabolic pathways, several of which have been described (Dobrosotskaya et al., 2002; Brice et al., 2009). The existence of direct biochemical connections between these pathways is further supported by the fact that the distribution of fatty acid chains between up- and downregulated species was similar for all lipid classes.

Surprisingly, the levels of dopamine and norepinephrine were altered in the striatum and adrenal gland, respectively, of Agpat4-deficient mice. In addition, we showed that the



Agpat4 gene promoter drove reporter gene expression in most catecholaminergic tissues. Altogether, these results suggest that Agpat4 may be involved in catecholamine metabolism. Although some lysophospholipid acyltransferases can process a broad range of substrates, including proteins (Zou et al., 2011), a direct effect of Agpat4 on catecholamines is improbable given the difference between catechol and glycerol-3-phosphate structures. Moreover, to our knowledge, no natural *O*-acyl derivatives of catecholamines have yet been isolated. Rather, an indirect relationship between glycerophospholipids and catecholamines should be considered. Acute intraventricular injection of lysophosphatidylcholine (LPC) has been shown to increase dopamine levels in the rat striatum, possibly through its action as a signaling molecule (Lee et al., 2005). LPC is produced by the hydrolysis of phosphatidylcholine, which is synthesized by the methylation of phosphatidylethanolamine. Interestingly, changes in the concentration of methylation markers have been observed in biological fluids of Parkinson's disease patients (Aoyama et al., 2000; Obeid et al., 2009). Therefore, although no differences in Agpat4 gene expression levels have been reported between Parkinson's disease patients and controls (Moran et al., 2006; Lesnick et al., 2007), our results raise the possibility that Agpat4 may play a role in the dopaminergic control of movement.

In this study, we report on the metabolic alterations caused by Agpat4 deficiency in mice. Broad perturbations in lysophosphatidylethanolamine levels suggest that Agpat4 may act as a lysophosphatidylethanolamine acyltransferase *in vivo*. Effects on catecholamine metabolism indicate a possible involvement in catecholamine-mediated functions, in particular in the midbrain dopaminergic structures, where Agpat4 expression is enriched. The Agpat4-deficient mouse line that we have generated constitutes a valuable tool to investigate these issues.

## Materials and methods

**Animals and tissue collection.** All experiments on mice were supervised by an investigator licensed by the Ministère de l'Agriculture, in conformity with French regulations.

Agpat4-deficient mice were generated from the Sanger Institute Gene Trap Resource ES cell line CC0173. The animals used in this study had a mixed background of ~ 6.25 % 129 and ~ 93.75 % C57BL/6. Of the 493 mice genotyped so far, 128 were Agpat4-deficient, 126 were wild-type and 239 were heterozygous.

Fresh tissues were collected immediately after euthanasia. Striatum and ventral midbrain were dissected from 1-mm-thick slices cut in a coronal matrix (Ted Pella) (Fig. S5). For Northern blot analysis, a 4-mm-thick block containing the entire ventral midbrain (between grooves 7 and 11 of the matrix) was removed from each brain.

**Genomic characterization of Agpat4-deficient mice.** The exact integration site of the gene-trap cassette within the first intron of the *Agpat4* gene was determined by Southern blot analysis and sequencing (Fig. S6). We also verified that the cassette integrated as a single copy (Fig. S7).

For genotyping, we developed a fast and convenient multiplex PCR assay using three primers (forward: 5' GGGCGCTCTGGGACTGA 3'; wild-type reverse: 5' ACATTACAGGACCCGTGTCTGA 3'; mutant reverse: 5' ACCAGACTCTCCCATCCACTACTC 3') and two TaqMan probes (Life Technologies) labeled with distinct dyes (wild-type probe: 6FAM-AAAAATTCACCATTATCCCACC; mutant probe: VIC-AACATGTGAGCCGAGGC). Amplification occurs from the forward and wild-type reverse primers for the wild-type allele, and from the forward and mutant reverse primers for the trapped allele, yielding, respectively, a 81-bp product detected by the

wild-type probe and a 108-bp product detected by the mutant probe. Each individual was genotyped in duplicate.

**Ultra-high performance liquid chromatography coupled with electrospray-LTQ-Orbitrap mass spectrometry (UPLC-MS).** Liquid-nitrogen-frozen ventral midbrains, striata and adrenal glands from individual mice were homogenized in 80 % methanol (500  $\mu$ l) using a Precellys-24 device (Bertin Technologies), then centrifuged twice at 18,000 x g for ten minutes. The supernatants were evaporated and the pellets resuspended in 100  $\mu$ l initial mobile phase, sonicated for ten minutes and centrifuged at 18,000 x g for ten minutes. Samples were randomized and quality controls containing internal standards interspersed. Ten microliters of each sample were injected onto a 2.1 x 150 mm Hypersil Gold C8 1.9  $\mu$ m column (Thermo Fisher Scientific) equipped with an on-line prefilter (Thermo Fisher Scientific). Mobile phases A and B were respectively 100 % water and 100 % acetonitrile, both containing 0.1 % formic acid. After an isocratic step of 2 min at 95 % phase A and 5 % phase B, a linear gradient from 5 to 100 % phase B was run over the next 11 min, followed by 12.5 min at 100 % phase B, before returning to 95 % phase A and 5 % phase B for 4.5 min, leading to a total run time of 30 min. The flow rate was 500  $\mu$ l/min.

The mass spectrometer was an LTQ-Orbitrap (Thermo Fisher Scientific) fitted with an electrospray source. Mass spectra were recorded from 60 to 2000 Th in both the positive and negative ion modes. In the positive ion mode, the electrospray voltage was set to 5 kV, the capillary voltage to 15 V, and the tube lens offset to 50 V. The sheath and auxiliary gas flows (both nitrogen) were respectively 60 and 10 (arbitrary units) and the drying gas temperature was 275°C. The same settings were used in the negative ion mode, except for the electrospray voltage, which was set to 3 kV, and the capillary voltage and tube lens offset, which were set to -10 V and -100 V, respectively. For collision-induced dissociation experiments, a

normalized collision energy ranging from 15 to 30 was used and the isolation width was set to 2 (arbitrary units).

**UPLC-MS data processing and analysis.** For each tissue, raw data were analyzed using the XCMS software (Smith et al., 2006). The resulting dataset consisted of a matrix with the m/z-retention time pairs as identifiers and their intensity values in each sample. To reduce background, we selected the m/z signals that were detected in at least five of the eight samples of either genotype and whose average intensity in the wild-type or knockout samples was greater than  $10^4$  (negative ion mode),  $5 \times 10^5$  (positive ion mode; first profiling experiment) or  $1 \times 10^5$  (positive ion mode; second profiling experiment), and greater than that in the blank samples ( $n = 8$ ). The curated matrices comprised between 2,036 and 8,112 m/z-retention time pairs, depending on the tissue and the experiment. These data were mean-centered and scaled to unit or Pareto variance, then introduced into SIMCA-P11 (Umetrics) for OPLS or PLS-DA analysis. Metabolites enriched in the wild-type or knockout tissues were selected among the m/z signals whose intensity was greater than  $10^5$  (negative ion mode),  $2.5 \times 10^6$  (positive ion mode; first profiling experiment) or  $5 \times 10^5$  (positive ion mode; second profiling experiment) in at least three samples of either genotype.

## Acknowledgments

The authors thank Nicolas Caudy (CEA, iBiTecS, SPI), Hélène Humbertclaude (CEA, iBiTecS, SBIGeM) and Sylvie Jounier (CEA, iBiTecS, SBIGeM) for technical assistance. They are grateful to Emmanuel Brouillet (CEA, I<sup>2</sup>BM, MIRCen, URA CEA-CNRS 2210) for fruitful discussions and friendly support.

N.L. was the recipient of a Ph.D. fellowship from the Ministère de l'Éducation Nationale, de la Recherche et de la Technologie. L.G. was supported by the Neuropôle de

Recherche Francilien and the Fondation pour la Recherche Médicale. We also received financial support from the IMAGEN European integrated project (EUR617037286).

## References

- Aoyama K, Matsubara K, Okada K, Fukushima S, Shimizu K, Yamaguchi S, Uezono T, Satomi M, Hayase N, Ohta S, Shiono H, Kobayashi S (2000). N-methylation ability for azaheterocyclic amines is higher in Parkinson's disease: nicotinamide loading test. *J Neural Transm* 107: 985-995.
- Beigneux AP, Vergnes L, Qiao X, Quatela S, Davis R, Watkins SM, Coleman RA, Walzem RL, Philips M, Reue K, Young SG (2006). Agpat6--a novel lipid biosynthetic gene required for triacylglycerol production in mammary epithelium. *J Lipid Res* 47: 734-744.
- Brice SE, Alford CW, Cowart LA (2009). Modulation of sphingolipid metabolism by the phosphatidylinositol-4-phosphate phosphatase Sac1p through regulation of phosphatidylinositol in *Saccharomyces cerevisiae*. *J Biol Chem* 284: 7588-7596.
- Brochier C, Gaillard MC, Diguët E, Caudy N, Dossat C, Ségurens B, Wincker P, Roze E, Caboche J, Hantraye P, Brouillet E, Elalouf JM, de Chaldée M (2008). Quantitative gene expression profiling of mouse brain regions reveals differential transcripts conserved in human and affected in disease models. *Physiol Genomics* 33: 170-179.
- Cao J, Li JL, Li D, Tobin JF, Gimeno RE (2006). Molecular identification of microsomal acyl-CoA:glycerol-3-phosphate acyltransferase, a key enzyme in de novo triacylglycerol synthesis. *Proc Natl Acad Sci U S A* 103: 19695-19700.
- Cao J, Shan D, Revett T, Li D, Wu L, Liu W, Tobin JF, Gimeno RE (2008). Molecular identification of a novel mammalian brain isoform of acyl-CoA:lysophospholipid acyltransferase with prominent ethanolamine lysophospholipid acylating activity, LPEAT2. *J Biol Chem* 283: 19049-19057.

- Dobrosotskaya IY, Seegmiller AC, Brown MS, Goldstein JL, Rawson RB (2002). Regulation of SREBP processing and membrane lipid production by phospholipids in *Drosophila*. *Science* 296: 879-883.
- Frisardi V, Panza F, Seripa D, Farooqui T, Farooqui AA (2011). Glycerophospholipids and glycerophospholipid-derived lipid mediators: A complex meshwork in Alzheimer's disease pathology. *Prog Lipid Res* [Epub ahead of print].
- Fujimoto H, Tadano-Aritomi K, Tokumasu A, Ito K, Hikita T, Suzuki K, Ishizuka I (2000). Requirement of seminolipid in spermatogenesis revealed by UDP-galactose: Ceramide galactosyltransferase-deficient mice. *J Biol Chem* 275: 22623-22626.
- Gadella BM, Colenbrander B, Van Golde LM, Lopes-Cardozo M (1993). Boar seminal vesicles secrete arylsulfatases into seminal plasma: evidence that desulfation of seminolipid occurs only after ejaculation. *Biol Reprod* 48: 483-489.
- Hishikawa D, Shindou H, Kobayashi S, Nakanishi H, Taguchi R, Shimizu T (2008). Discovery of a lysophospholipid acyltransferase family essential for membrane asymmetry and diversity. *Proc Natl Acad Sci U S A* 105: 2830-2835.
- Honke K, Yamane M, Ishii A, Kobayashi T, Makita A (1996). Purification and characterization of 3'-phosphoadenosine-5'-phosphosulfate:GalCer sulfotransferase from human renal cancer cells. *J Biochem* 119: 421-427.
- Hughey CA, Hendrickson CL, Rodgers RP, Marshall AG, Qian K (2001). Kendrick mass defect spectrum: a compact visual analysis for ultrahigh-resolution broadband mass spectra. *Anal Chem* 73: 4676-4681.
- Kopp F, Komatsu T, Nomura DK, Trauger SA, Thomas JR, Siuzdak G, Simon GM, Cravatt BF (2010). The glycerophospho metabolome and its influence on amino acid homeostasis revealed by brain metabolomics of GDE1(-/-) mice. *Chem Biol* 17: 831-840.

- Lands WE (2000). Stories about acyl chains. *Biochim Biophys Acta* 1483: 1-14.
- Lee ES, Soliman KF, Charlton CG (2005). Lysophosphatidylcholine decreases locomotor activities and dopamine turnover rate in rats. *Neurotoxicology* 26: 27-38.
- Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, Boe AF, Boguski MS, Brockway KS, Byrnes EJ, Chen L, Chen L, Chen TM, Chin MC, Chong J, Crook BE, Czaplinska A, Dang CN, Datta S, Dee NR, Desaki AL, Desta T, Diep E, Dolbeare TA, Donelan MJ, Dong HW, Dougherty JG, Duncan BJ, Ebbert AJ, Eichele G, Estin LK, Faber C, Facer BA, Fields R, Fischer SR, Fliss TP, Frensley C, Gates SN, Glattfelder KJ, Halverson KR, Hart MR, Hohmann JG, Howell MP, Jeung DP, Johnson RA, Karr PT, Kawal R, Kidney JM, Knapik RH, Kuan CL, Lake JH, Laramie AR, Larsen KD, Lau C, Lemon TA, Liang AJ, Liu Y, Luong LT, Michaels J, Morgan JJ, Morgan RJ, Mortrud MT, Mosqueda NF, Ng LL, Ng R, Orta GJ, Overly CC, Pak TH, Parry SE, Pathak SD, Pearson OC, Puchalski RB, Riley ZL, Rockett HR, Rowland SA, Royall JJ, Ruiz MJ, Sarno NR, Schaffnit K, Shapovalova NV, Sivisay T, Slaughterbeck CR, Smith SC, Smith KA, Smith BI, Sodt AJ, Stewart NN, Stumpf KR, Sunkin SM, Sutram M, Tam A, Teemer CD, Thaller C, Thompson CL, Varnam LR, Visel A, Whitlock RM, Wohnoutka PE, Wolkey CK, Wong VY, Wood M, Yaylaoglu MB, Young RC, Youngstrom BL, Yuan XF, Zhang B, Zwingman TA, Jones AR (2007). Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445: 168-176.
- Lesnick TG, Papapetropoulos S, Mash DC, Ffrench-Mullen J, Shehadeh L, de Andrade M, Henley JR, Rocca WA, Ahlskog JE, Maraganore DM (2007). A genomic pathway approach to a complex disease: axon guidance and Parkinson disease. *PLoS Genet* 3: e98.



- Lu B, Jiang YJ, Zhou Y, Xu FY, Hatch GM, Choy PC (2005a). Cloning and characterization of murine 1-acyl-sn-glycerol 3-phosphate acyltransferases and their regulation by PPARalpha in murine heart. *Biochem J* 385: 469-477.
- Lu B, Jiang YJ, Man MQ, Brown B, Elias PM, Feingold KR (2005b). Expression and regulation of 1-acyl-sn-glycerol- 3-phosphate acyltransferases in the epidermis. *J Lipid Res* 46: 2448-2457.
- Moran LB, Duke DC, Deprez M, Dexter DT, Pearce RK, Graeber MB (2006). Whole genome expression profiling of the medial and lateral substantia nigra in Parkinson's disease. *Neurogenetics* 7: 1-11.
- Nagle CA, Vergnes L, Dejong H, Wang S, Lewin TM, Reue K, Coleman RA (2008). Identification of a novel sn-glycerol-3-phosphate acyltransferase isoform, GPAT4, as the enzyme deficient in *Agpat6*<sup>-/-</sup> mice. *J Lipid Res* 49: 823-831.
- Obeid R, Schadt A, Dillmann U, Kostopoulos P, Fassbender K, Herrmann W (2009). Methylation status and neurodegenerative markers in Parkinson disease. *Clin Chem* 55: 1852-1860.
- Påhlsson P, Spitalnik SL, Spitalnik PF, Fantini J, Rakotonirainy O, Ghardashkhani S, Lindberg J, Konradsson P, Larson G (2001). Characterization of galactosyl glycerolipids in the HT29 human colon carcinoma cell line. *Arch Biochem Biophys* 396: 187-198.
- Pete MJ, Ross AH, Exton JH (1994). Purification and properties of phospholipase A1 from bovine brain. *J Biol Chem* 269: 19494-19500.
- Pieringer J, Rao GS, Mandel P, Pieringer RA (1977). The association of the sulphogalactosylglycerolipid of rat brain with myelination. *Biochem J* 166: 421-428.
- Prasad SS, Garg A, Agarwal AK (2011). Enzymatic activities of the human AGPAT isoform 3 and isoform 5: localization of AGPAT5 to mitochondria. *J Lipid Res* 52: 451-462.

- Saghatelian A, Trauger SA, Want EJ, Hawkins EG, Siuzdak G, Cravatt BF (2004). Assignment of endogenous substrates to enzymes by global metabolite profiling. *Biochemistry* 43: 14332-14339.
- Shindou H, Shimizu T (2009). Acyl-CoA:lysophospholipid acyltransferases. *J Biol Chem* 284: 1-5.
- Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G (2006). XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* 78: 779-787.
- Vergnes L, Beigneux AP, Davis R, Watkins SM, Young SG, Reue K (2006). Agpat6 deficiency causes subdermal lipodystrophy and resistance to obesity. *J Lipid Res* 47: 745-754.
- Yuki K, Shindou H, Hishikawa D, Shimizu T (2009). Characterization of mouse lysophosphatidic acid acyltransferase 3: an enzyme with dual functions in the testis. *J Lipid Res* 50: 860-869.
- Zhao Y, Chen YQ, Li S, Konrad RJ, Cao G (2009). The microsomal cardiolipin remodeling enzyme acyl-CoA lysocardiolipin acyltransferase is an acyltransferase of multiple anionic lysophospholipids. *J Lipid Res* 50: 945-956.
- Zou C, Ellis BM, Smith RM, Chen BB, Zhao Y, Mallampalli RK (2011). Acyl-CoA:lysophosphatidylcholine acyltransferase I (Lpcat1) catalyzes histone protein O-palmitoylation to regulate mRNA synthesis. *J Biol Chem* [Epub ahead of print].

## Supplementary Data

**Table S1. Probes and primers**

|                 | Gene    | Description  |
|-----------------|---------|--|
| Southern probes | Agpat4  | A 276-bp fragment of the Agpat4 gene was amplified by PCR using primers (5' TTAACAATCGGGCTATGCCTT 3') and (5' GACTTCCACGGGAGTGCAAA 3') and subcloned into pCR4Blunt-TOPO (Life Technologies). The resulting construct was digested with <i>EcoR</i> I, yielding a 294-bp probe.  |
|                 | β-gal   | A 827-bp <i>EcoR</i> V- <i>Sac</i> I fragment of the beta-galactosidase transgene was used as a probe.   |
|                 | Capucin | A 1,123-bp fragment of the Capucin gene was amplified by PCR using primers (5' CCCAAATCTCTTCTTACCCTAGA 3') and (5' ACTGTGATGCGAATGTACCC 3') and subcloned into pCR4Blunt-TOPO (Life Technologies). The resulting construct was digested with <i>EcoR</i> I and <i>Bgl</i> I, yielding a 882-bp probe.                      |
| Northern probe  | Agpat4  | A 1,161-bp fragment encompassing the entire coding sequence of Agpat4 was amplified by PCR using primers (5' GATTTATCTCTTGAGAATCCCCACACC 3') and (5' GTCCGTTTGTTCGGTTTGTGTC 3') and subcloned into pCR4Blunt-TOPO (Life Technologies). The resulting construct was digested with <i>EcoR</i> I, yielding a 1,179-bp probe. |
| qPCR primers    | Agpat4  | 5' CGTGTTTCATCGCCTCGG 3'<br>5' GCAGAGTCTGGCATTGATCTTG 3'<br>These primers amplify a fragment of <i>Agpat4</i> exon 2.  |
|                 | Agpat3  | 5' TCTTCGTGGTGAGCGGGTTA 3'<br>5' CAAGCGGCAGTTGATACGG 3'  |
|                 | Lclat1  | 5' ACTTCCACGTCCAGCGGTAT 3'<br>5' CTTGGTAGAAGGACCGAAGCC 3'  |
|                 | Agpat5  | 5' TGGCTTCCGCTGTATGGGT 3'<br>5' CATTGGTGTTCTGCGTTCAC 3'  |
|                 | Agpat1  | 5' AGGTCCTGCCAGATCGTTGT 3'<br>5' GTGCGTTTCCGGTCAATG 3'   |
|                 | Agpat2  | 5' TGTCTGTGATGGCCGACCTG 3'<br>5' CCCATTGTGCTTGCGTGTACC 3'  |
|                 | Lpcat4  | 5' GCCTTTGCCCAACAGTTACAG 3'<br>5' GCCACATTTGGAAGTCAACC 3'  |
|                 | Ppia    | 5' ATGGCAAATGCTGGACCAAA 3'<br>5' GCCTTCTTTCACCTTCCCAAA 3'  |

All probes were random prime-labeled with [ $\alpha$ -<sup>32</sup>P]dCTP. Peptidylpropyl isomerase A (*Ppia*) was used as an internal control for real-time quantitative PCR (qPCR) experiments.

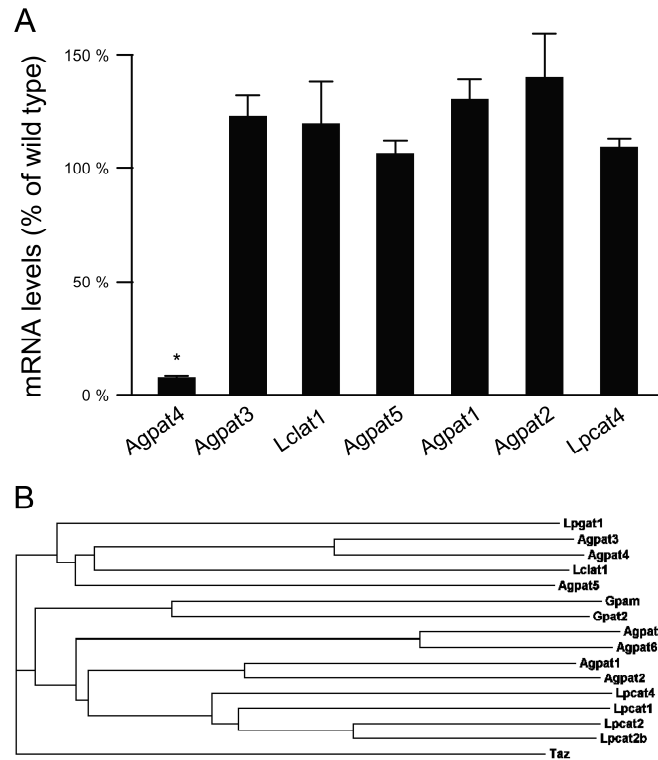


Fig. S1

**Fig. S1.** Expression levels of Agpat family members in Agpat4-deficient mice. (A) mRNA levels were measured in the ventral midbrain of seven Agpat4-deficient mice and seven wild-type littermates by real-time quantitative PCR. Primer sequences are listed in Table S1. Bars represent the average mRNA levels in Agpat4-deficient samples relative to wild-type controls  $\pm$  SEM. Significant differences were assessed using unpaired Student's t test. \*  $P = 1.10^{-9}$ . (B) Phylogram of mouse Agpat family members generated from a ClustalW2 (European Molecular Biology Laboratory-European Bioinformatics Institute) protein alignment.

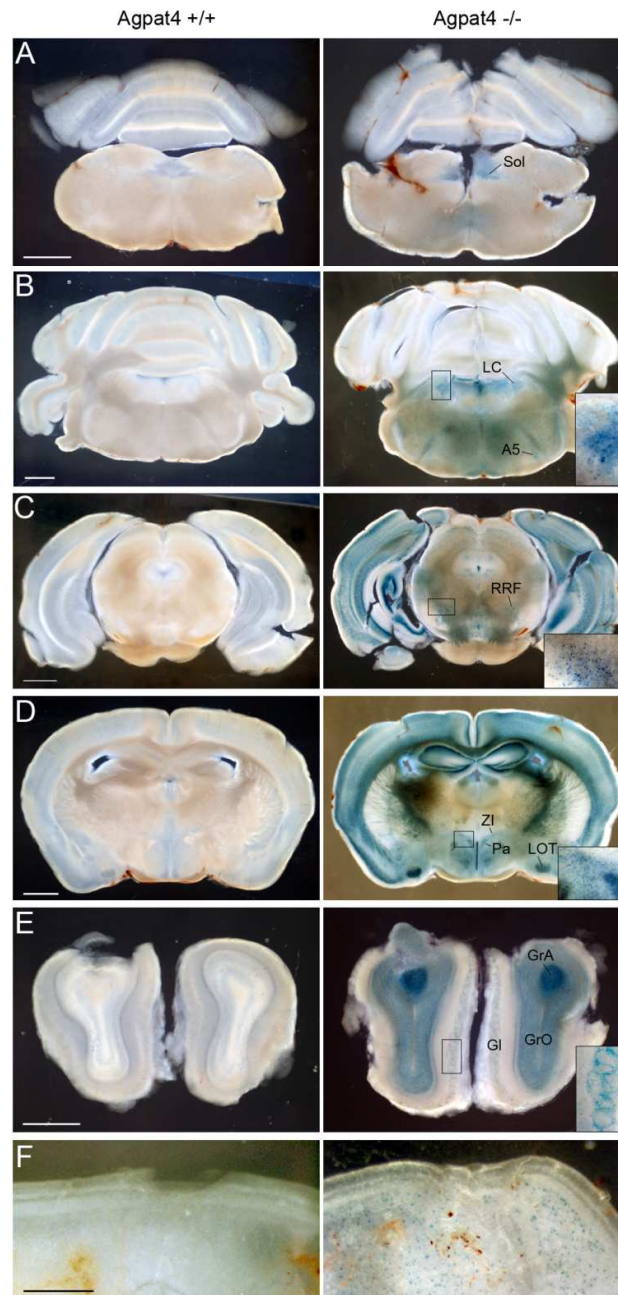


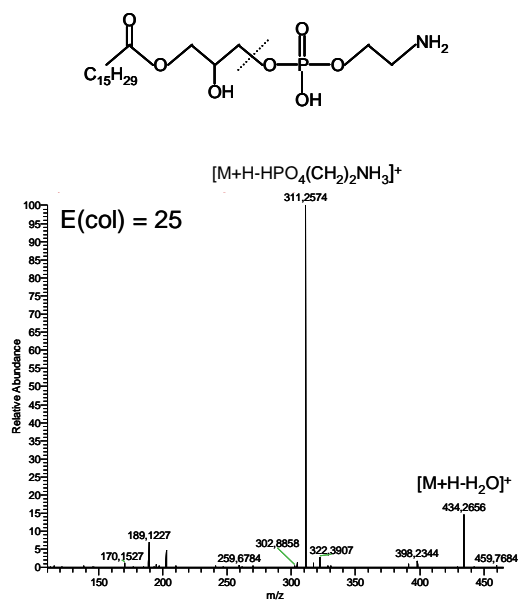
Fig. S2

**Fig. S2.** X-Gal staining of 300- $\mu$ m coronal brain sections (A-E) and 100- $\mu$ m retinal sections (F) from wild-type and Agpat4-deficient littermates. Stained noradrenergic structures include: (A) the solitary-vagal complex (Sol); (B) the locus coeruleus (LC) and A5 cell group. Stained regions containing dopaminergic cells include: (C) the retrorubral field (RRF); (D) the zona incerta (ZI) and paraventricular nucleus (Pa); (E) the shell of the olfactory glomeruli (Gl); (F) the retina. Note the intense staining of the nucleus of the lateral olfactory tract (LOT) and the granular cell layers of the main (GrO) and accessory (GrA) olfactory bulbs. The anterior olfactory nucleus is also strongly stained (not shown). (Scale bars: 1 mm for brain sections; 0.5 mm for retinal sections.)

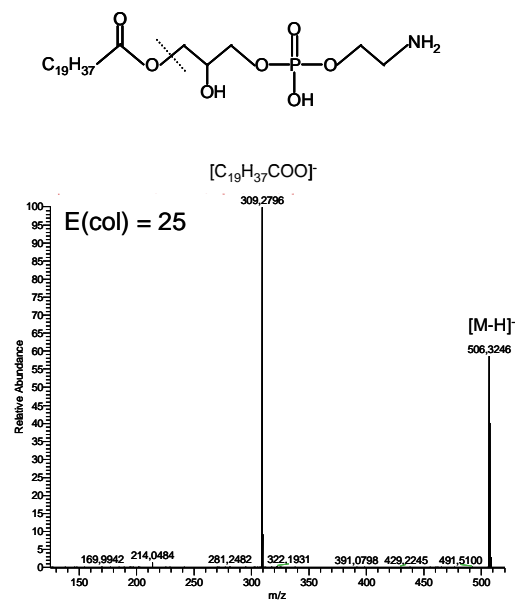
**Table S2. Kendrick mass defect analysis of ventral midbrain lipids**

| Metabolites                   | m/z [M-H] <sup>-</sup> | Retention time  | Exact mass      | Kendrick mass   | Kendrick nominal mass | Kendrick mass defect |
|-------------------------------|------------------------|-----------------|-----------------|-----------------|-----------------------|----------------------|
| Lysophosphatidylethanolamines |                        |                 |                 |                 |                       |                      |
| <b>LPE(14:0)</b>              | <b>424.2437</b>        | <b>10.4 min</b> | <b>425.2510</b> | <b>424.7761</b> | <b>424</b>            | <b>0.78</b>          |
| LPE(15:0)                     | 438.2608               | 10.9 min        | 439.2681        | 438.7776        | 438                   | 0.78                 |
| <b>LPE(16:0)</b>              | <b>452.2772</b>        | <b>11.4 min</b> | <b>453.2844</b> | <b>452.7783</b> | <b>452</b>            | <b>0.78</b>          |
| LPE(18:0)                     | 480.3071               | 12.7 min        | 481.3144        | 480.7770        | 480                   | 0.78                 |
| <b>LPE(16:1)</b>              | <b>450.2595</b>        | <b>10.7 min</b> | <b>451.2667</b> | <b>450.7628</b> | <b>450</b>            | <b>0.76</b>          |
| LPE(17:1)                     | 464.2761               | 11.2 min        | 465.2834        | 464.7638        | 464                   | 0.76                 |
| LPE(18:1)                     | 478.2928               | 11.7 min        | 479.3001        | 478.7649        | 478                   | 0.76                 |
| LPE(19:1)                     | 492.3061               | 12.2 min        | 493.3133        | 492.7625        | 492                   | 0.76                 |
| <b>LPE(20:1)</b>              | <b>506.3223</b>        | <b>12.7 min</b> | <b>507.3296</b> | <b>506.7631</b> | <b>506</b>            | <b>0.76</b>          |
| LPE(18:2)                     | 476.2754               | 11.1 min        | 477.2827        | 476.7498        | 476                   | 0.75                 |
| <b>LPE(20:2)</b>              | <b>504.3074</b>        | <b>12.1 min</b> | <b>505.3147</b> | <b>504.7505</b> | <b>504</b>            | <b>0.75</b>          |
| <b>LPE(20:3)</b>              | <b>502.2910</b>        | <b>11.5 min</b> | <b>503.2983</b> | <b>502.7363</b> | <b>502</b>            | <b>0.74</b>          |
| LPE(22:3)                     | 530.3224               | 12.6 min        | 531.3297        | 530.7364        | 530                   | 0.74                 |
| LPE(20:4)                     | 500.2765               | 11.1 min        | 501.2838        | 500.7240        | 500                   | 0.72                 |
| <b>LPE(22:4)</b>              | <b>528.3068</b>        | <b>11.9 min</b> | <b>529.3141</b> | <b>528.7231</b> | <b>528</b>            | <b>0.72</b>          |
| LPE(24:4)                     | 556.3379               | 12.7 min        | 557.3452        | 556.7229        | 556                   | 0.72                 |
| Monogalactosylglycerols       |                        |                 |                 |                 |                       |                      |
| (0:0/16:0) + FA               | 537.3272               | 12.2 min        | 538.3345        | 537.7334        | 537                   | 0.73                 |
| <b>(O-16:0/16:0) + FA</b>     | <b>761.5729</b>        | <b>21.2 min</b> | <b>762.5801</b> | <b>761.7286</b> | <b>761</b>            | <b>0.73</b>          |

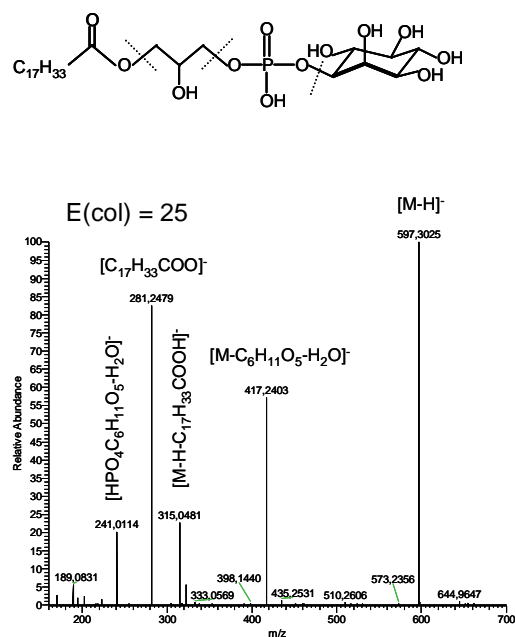
M/z data and retention times are average values from eight Agpat4-deficient and eight wild-type samples. Metabolites in bold have been analyzed by MS/MS (Fig. S3).



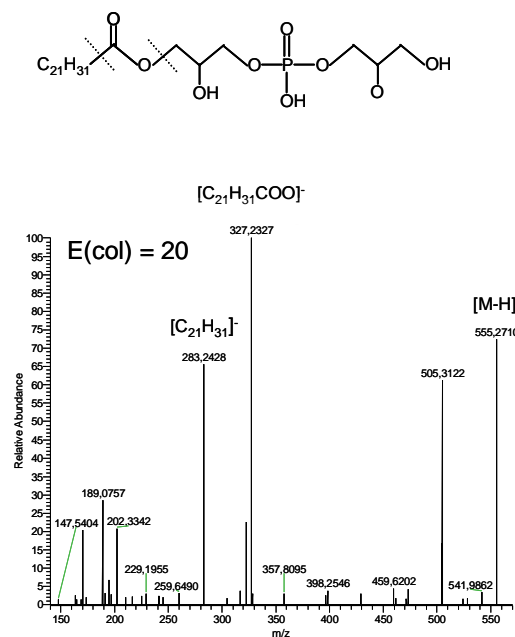
LPE(16:1)



LPE(20:1)



LPI(18:1)



LPG(22:6)

Fig. S3 (page 1/2)

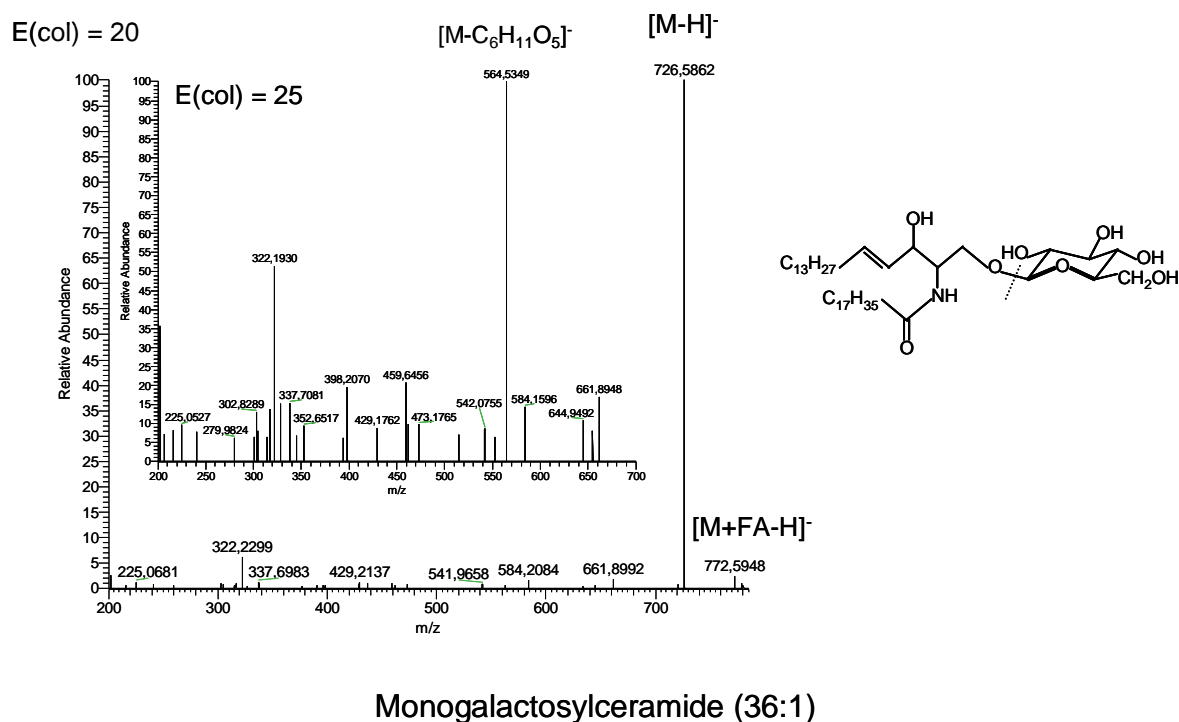
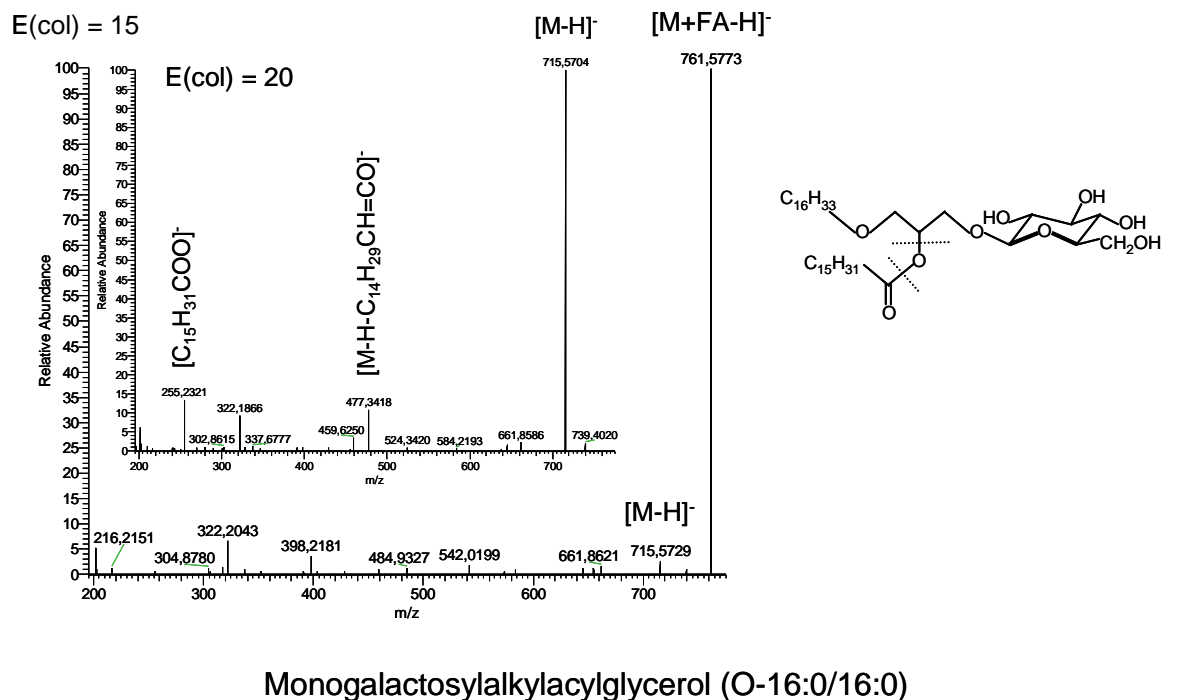


Fig. S3 (page 2/2)

**Fig. S3.** Examples of MS/MS spectra. A structure is proposed for each example. The fatty acid chain of lysophospholipids has been arbitrarily placed in the *sn*-1 position of the glycerol backbone. Dotted lines indicate the fragmentation sites.  $E(\text{col})$ , normalized collision energy; FA, formic acid.



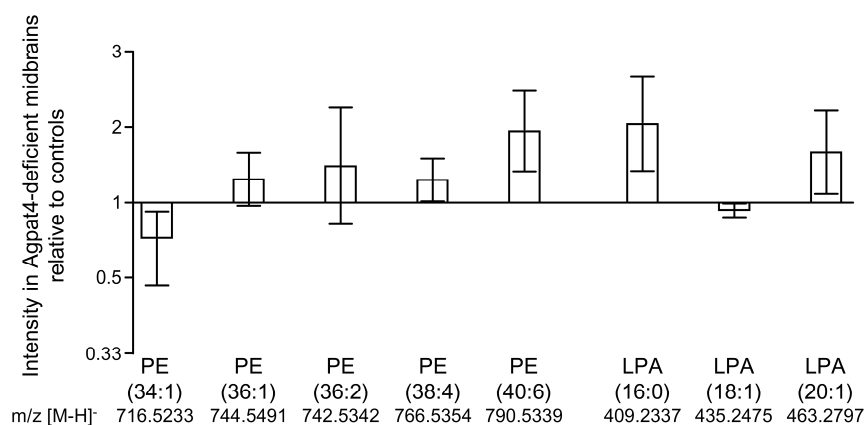


Fig. S4

**Fig. S4.** Levels of putative diacylphosphatidylethanolamines and phosphatidic acids in the ventral midbrain of *Agpat4*-deficient mice. Bars represent the average ratio  $\pm$  SEM of the signal intensity in *Agpat4*-deficient mice ( $n = 8$ ) to the mean signal intensity in wild-type littermates ( $n = 8$ ). Results are shown for one profiling experiment for  $[M-H]^-$  signals whose intensity was greater than  $10^5$  in at least three samples of either genotype. PE, phosphatidylethanolamine; LPA, lysophosphatidic acid.

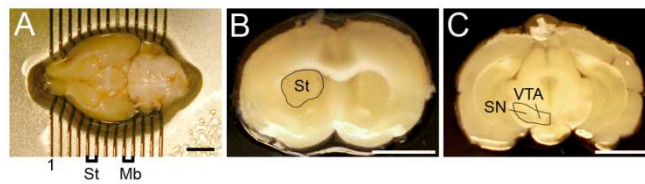


Fig. S5

**Fig. S5.** Dissection of brain regions. (A) Fresh brains were placed ventral side up in an ice-cold coronal matrix fitted with grooves spaced one millimeter apart. The rostral-most groove is indicated by number “1” in the figure. Brain slices were cut by inserting razor blades into the grooves, then transferred rostral side up into drops of PBS in an ice-cold dish. The striatum (St) and ventral midbrain (Mb) were dissected bilaterally using a scalpel as depicted in (B) and (C), respectively. The ventral midbrain contains the substantia nigra (SN) and the ventral tegmental area (VTA). (Scale bars, 3 mm.)

A

gctcagtgatgtagggagagcaggggttatcattgaagaccatgagcatcctgtctggaaggacggcgcatctggg  
12,335,830  
↓  
gcgctctgggactgagcgggcaagggGATAACGCAGGAAAGAACATGTGAGCCGAGGCTTGCCGCCCCTGACTGCA  
CTGCACTGAGTAGTGGATGGGAGAGTCTGGTACCGCACTGCCGTTTCCTCCACCATCCCCGCAGCGCAGGGCA...

B

...AATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGTATTGGGCGCTCTTCGCTTCCTCGCTCACTGACTCGCTGC  
12,335,854  
↓  
GCTCGGTCGTTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAatcccaccctcagacacgggtcctg  
taatgttgaaggttggtcatggtgggaggggtcccgctctgggtcctctttatttacaggagtggttggtgcag

Fig. S6

**Fig. S6.** Integration site of the gene-trap cassette within the first intron of the *Agpat4* gene. (A) Cassette 5'-junction. (B) Cassette 3'-junction. The intron sequence is in lower case; the cassette sequence is in capital letters. Numbers refer to nucleotide coordinates on mouse chromosome 17.

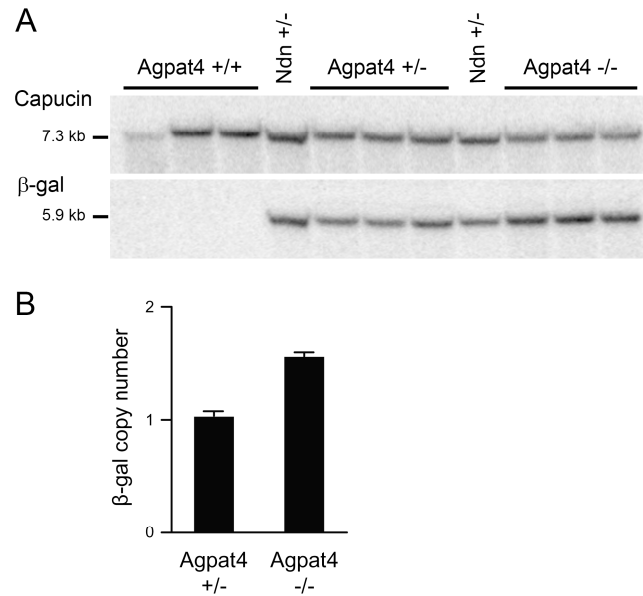


Fig. S7

**Fig. S7.** Determination of the gene-trap cassette copy number in *Agpat4*-deficient mice. (A) Southern blot analysis of *EcoRI*-digested genomic DNA from different mice using a probe for the beta-galactosidase transgene ( $\beta$ -gal) and a probe for the Capucin gene to normalize for DNA input. Probe information is provided in Table S1. Heterozygous Necdin-mutant mice (*Ndn* +/-) have a single  $\beta$ -gal copy in their genome (Gérard M, Hernandez L, Wevrick R, Stewart CL [1999]. Disruption of the mouse necdin gene results in early post-natal lethality. *Nat Genet* 23: 199-202). (B) For each sample, the  $\beta$ -gal signal intensity, as determined by ImageJ software (National Institute of Health), was divided by the Capucin signal intensity for normalization. Bars represent the average ratio  $\pm$  SEM of the normalized  $\beta$ -gal signal intensities for the *Agpat4* +/- or *Agpat4* -/- samples ( $n = 3$ ) to the mean normalized  $\beta$ -gal signal intensity for the *Ndn* +/- samples.

## 4. Discussion

Au cours de ce travail j'ai analysé les propriétés d'une lignée de souris déficientes pour le gène d'Agpat4. Ces souris sont apparemment en bonne santé, fertiles, ne présentent pas de mortalité précoce ni d'anomalie comportementale évidente. Il arrive parfois que l'inactivation d'un gène par la technique de « *gene trap* » ne soit pas complète (Voss et al., 1998). Nous avons voulu de ce fait vérifier l'efficacité de l'invalidation du gène Agpat4 dans notre lignée de souris. Le résultat d'une analyse par Northern blot montre que le niveau d'expression d'Agpat4 est considérablement plus faible dans le cerveau des souris transgéniques que chez les souris sauvages. L'absence de phénotype évident pourrait aussi s'expliquer par une éventuelle compensation du défaut d'Agpat4 chez les souris transgéniques par d'autres membres de la famille Agpat. Nous avons donc évalué par RT-PCR quantitative les taux d'expression de plusieurs d'entre eux. Cette analyse ne montre pas de différences significatives d'expression de ces homologues entre souris transgéniques et sauvages. Ceci suggère soit que le défaut d'expression d'Agpat4 est compensé par la surexpression d'autres enzymes dont nous n'avons pas testé le niveau d'expression, soit que les niveaux d'expression sauvages des enzymes capables de catalyser la ou les même(s) réaction(s) qu'Agpat4 sont suffisants pour compenser son défaut. En effet, il existe une grande redondance fonctionnelle parmi les Agpats (Hishikawa et al., 2008 ; Shindou et Shimizu 2009).

Le gène Agpat4 présente un profil d'expression original parmi les autres Agpat. En particulier, il est peu exprimé dans les tissus adipeux blanc et brun (Vergnes et al., 2006). Dans cette étude, nous avons analysé son profil d'expression grâce au gène rapporteur codant la  $\beta$ -galactosidase, présent dans la cassette de « *gene trap* », ainsi que chez les souris sauvage par RT-PCR quantitative. Les résultats montrent une expression dans plusieurs régions du cerveau (notamment l'hippocampe), et tissus périphériques (notamment l'épiderme, les testicules, les poumons et différents étages du tube digestif, en particulier le côlon). Une expression enrichie dans le cerveau, le côlon, les testicules, les poumons et la peau a déjà été observée auparavant (Lu et al., 2005 (1) ; Lu et al., 2005 (2) ; Beigneux et al., 2006). En revanche l'expression d'Agpat4 dans la plupart des tissus catécholaminergiques aussi bien dans le cerveau (la substance noire, l'aire tegmentaire ventrale, le locus coeruleus, le noyau solitaire) qu'en périphérie (les ganglions cervicaux supérieurs, la médulla des glandes surrénales) est un résultat nouveau.

Agpat4 étant une acyltransférase putative, nous avons comparé les métabolomes des glandes surrénales, du mésencéphale ventral, tissus dans lesquels Agpat4 est exprimé, et du striatum, qui contient les projections dopaminergiques de la substance noire, entre les souris sauvages et transgéniques. Nous avons constaté une perturbation du métabolisme des glycérophospholipides et plus particulièrement de celui des lysophosphatidyléthanolamines dans tous les tissus analysés. L'expression préférentielle d'Agpat4 dans le cerveau, un tissu particulièrement riche en phosphatidyléthanolamines, suggère que la principale fonction physiologique d'Agpat4 puisse être le remodelage des phosphatidyléthanolamines. Chez les souris transgéniques nous avons observé que les lysophosphatidyléthanolamines enrichies dans les tissus transgéniques avaient de courtes chaînes d'acide gras portant peu d'insaturations, tandis que les lysophosphatidyléthanolamines appauvries avaient des chaînes d'acide gras plus longues avec plus d'insaturations. Les chaînes d'acides gras localisées en position *sn2* des glycérophospholipides sont généralement plus longues et plus insaturées que celles situées en position *sn1*. Ceci suggère qu'Agpat4 pourrait catalyser une acylation en position *sn2* des lysophosphatidyléthanolamines. Parmi les Agpats connues, Agpat7 (Cao et al., 2008) et Agpat5 (Prasad et al., 2011) possèdent une activité lysophosphatidyléthanolamine acyltransférase. La compensation éventuelle de l'expression d'Agpat4 chez les souris transgéniques par surexpression des gènes de ces enzymes a été analysée par RT-PCR quantitative, mais nous n'avons pas vu de différence significative entre les souris sauvages et transgéniques. Cependant il existe d'autres lysophosphatidyléthanolamine acyltransférases (MBOAT1, MBOAT2, LPCAT3 (Hishikawa et al., 2008)) dont l'effet compensateur éventuel pourrait être testé. Cette fonction hypothétique pourrait être confirmée par des expériences *in vitro*. Pour ce faire, la protéine Agpat4 recombinante étiquetée pourrait être surexprimée dans des cellules de mammifères. Etant donné qu'Agpat4 est une protéine membranaire ce système paraît le plus approprié. Il est d'ailleurs largement utilisé pour des études similaires sur d'autres membres de la famille Agpat (Prasad et al., 2011 ; Agarwal et al., 2007 ; Cao et al., 2008 ; Eberhardt et al., 1997). L'activité enzymatique serait ensuite testée en utilisant un extrait protéique des cellules transfectées, en présence de lysophosphatidyléthanolamine et d'Acyl-coA synthétiques. Les métabolites pourraient être détectés par spectrométrie de masse couplée à la chromatographie liquide.

Les phosphatidyléthanolamines sont enrichies dans le cerveau. Dans l'hippocampe elles représentent environ 45% de la totalité des phospholipides (Delion et al., 1997). Leur fonction a été étudiée. Notamment la 1,2-dilinoyleoyl-*sn*-glycero-3-phosphoéthanolamine

améliore la mémoire spatiale, probablement par une action protectrice contre la mort neuronale (Yaguchi et al., 2010). L'influence du vieillissement sur la composition lipidique de l'hippocampe a été étudiée chez des rats. Une baisse des phospholipides majeurs (phosphatidyléthanolamines, phosphatidylcholines) a été constatée tandis que le niveau de phospholipides mineurs (phosphatidylserines, phosphatidylinositols et sphingomyélines) est augmenté (Delion et al., 1997). Le changement de la composition lipidique de l'hippocampe induit par l'âge pourrait avoir un effet sur son fonctionnement, notamment sur la mémoire. Agpat4, qui est exprimée au niveau de l'hippocampe, pourrait donc avoir un rôle dans le fonctionnement de la mémoire via la production de phosphatidyléthanolamines. Les niveaux de phosphatidyléthanolamines baissent également chez les patients de la maladie d'Alzheimer (Nitsch et al., 1992), ainsi que de la maladie de Parkinson (Riekkinen et al., 1975), ce qui suggère une augmentation de la dégradation des phospholipides membranaires dans le cerveau au cours des maladies neurodégénératives. Pour évaluer le rôle potentiel d'Agpat4 dans la mémoire, des tests comportementaux chez les souris transgéniques pourraient être envisagés, par exemple le test de la piscine de Morris et le test du « *Y-maze* ».

D'autres classes de lipides perturbés chez les souris transgéniques ont été identifiées dans cette étude, notamment des glycolipides (deux galactosylglycérols et un cérébroside potentiel). La distribution des chaînes d'acides gras entre les glycolipides enrichis et diminués chez les souris transgéniques est similaire à celle observée pour les lysophosphatidyléthanolamines. Ce résultat suggère qu'il y a probablement une connexion entre le métabolisme des glycolipides et celui des glycérophospholipides. Cette connexion pourrait se faire par un ou des substrat(s) commun(s) aux deux voies métaboliques. Les galactosylglycérols et les cérébrosides font partie de la composition lipidique de la myéline. Il a été démontré sur l'exemple d'une leukodystrophie que la perturbation du métabolisme lipidique peut provoquer une production de myéline pathologique (Ginsberg et Gershfeld, 1991). L'impact d'Agpat4 sur la composition de myéline et donc sur ses propriétés pourrait être analysé chez les souris transgéniques à plusieurs échelles : histologique, biochimique (analyse de la composition de la myéline par spectrométrie de masse), mais aussi à l'échelle comportementale (recherche d'un phénotype neurologique).

Une perturbation du métabolisme des catécholamines a également été constatée chez les souris transgéniques. Cette observation suggère une connexion inattendue entre le métabolisme des glycérophospholipides et celui des catécholamines. Il s'agit vraisemblablement d'un mécanisme indirect, car aucun dérivé O-acyl naturel des

catécholamines n'a encore été trouvé. Une action indirecte pourrait s'expliquer par le rôle des glycérophospholipides dans la signalisation cellulaire. En effet les lysophospholipides ont des fonctions variées dans les processus cellulaires notamment comme médiateurs via leur récepteurs couplés aux protéines G (Makide et al., 2009). Une autre explication possible de ce résultat pourrait être l'action directe des lipides membranaires sur les récepteurs dopaminergiques.

L'expression d'Agpat4 est enrichie dans les principales structures dopaminergiques du cerveau. Le phénotype dopaminergique des souris transgénique pourrait ainsi être examiné. Différentes approches sont envisageables : une approche histologique (comptage des neurones dopaminergiques, analyse morphologique, marquage contre l'ubiquitine et l' $\alpha$ -synucléine pour visualiser d'éventuelles inclusions) ; une approche comportementale (analyse du comportement locomoteur (test du rotarod, « *pole test* », activité spontanée en *openfield*) ; une approche pharmacologique (analyse de l'effet de différents agents pharmacologiques comme l'amphétamine, qui inhibe la recapture de la dopamine, ou l'apomorphine, qui est un agoniste des récepteurs dopaminergiques D1 et D2, sur le comportement locomoteur des souris transgéniques). On pourrait aussi analyser ces souris transgéniques dans le contexte de la maladie de Parkinson (par exemple en les traitant au MPTP, qui reproduit un phénotype Parkinsonien). Les mêmes analyses histologiques et comportementales que celles décrits ci-dessus pourraient être réalisées sur les souris transgéniques traitées.



## BIBLIOGRAPHIE

- Agarwal A.K., Arioglu E., De Almeida S., Akkoc N., Taylor S.I., Bowcock A.M., Barnes R.I., and Garg A. (2002) AGPAT2 is mutated in congenital generalized lipodystrophy linked to chromosome 9q34. *Nat. Genet.* **31**:21-23.
- Agarwal A.K., Barnes R.I., and Garg A. (2006) Functional characterization of human 1-acylglycerol-3-phosphate acyltransferase isoform 8: cloning, tissue distribution, gene structure, and enzymatic activity. *Arch. Biochem. Biophys.* **449**:64-76.
- Agarwal A.K., and Garg A. (2003) Congenital generalized lipodystrophy: significance of triglyceride biosynthetic pathways. *Trends Endocrinol. Metab.* **14**:214-221.
- Agarwal A.K., and Garg A. (2010) Enzymatic activity of the human 1-acylglycerol-3-phosphate-O-acyltransferase isoform 11: upregulated in breast and cervical cancers. *J. Lipid Res.* **51**:2143-52.
- Agarwal A.K., Sukumaran S., Bartz R., Barnes R.I., and Garg A. (2007) Functional characterization of human 1-acylglycerol-3-phosphate-O-acyltransferase isoform 9: cloning, tissue distribution, gene structure, and enzymatic activity. *J. Endocrinol.* **193**:445-457.
- Agaton C., Galli J., Höiden Guthenberg I., Janzon L., Hansson M., Asplund A., Brundell E., Lindberg S., Ruthberg I., Wester K., Wurtz D., Höög C., Lundeberg J., Ståhl S., Pontén F., and Uhlén M. (2003) Affinity proteomics for systematic protein profiling of chromosome 21 gene products in human tissues. *Mol. Cell. Proteomics* **2**:405-14.
- Aguado B., and Campbell R.D. (1998) Characterization of a human lysophosphatidic acid acyltransferase that is encoded by a gene located in the class III region of the human major histocompatibility complex. *J. Biol. Chem.* **273**:4096-105.
- Albin R.L., Young A.B., and Penney J.B. (1989) The functional anatomy of basal ganglia disorders. *Trends. Neurosci.* **12**: 366-75.
- Alexander G.E., DeLong M.R., and Strick P.L. (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* **9**: 357-81.
- Allinson J., and Brooks S. (2004) Biomarkers in Drug Development – A CRO Perspective *Current Separations* **21**: 1.
- Altschul S.F., Gish W., Miller W., Myers E.W., and Lipman D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**:403-10.

- Aoyama K., Suh S.W., Hamby A.M., Liu J., Chan W.Y., Chen Y., and Swanson R.A. (2006) Neuronal glutathione deficiency and age-dependent neurodegeneration in the EAAC1 deficient mouse. *Nat. Neurosci.* **9**:119-126.
- Araki Y., Yoshikawa K., Okamoto S., Sumitomo M., Maruwaka M., and Wakabayashi T. (2010) Identification of novel biomarker candidates by proteomic analysis of cerebrospinal fluid from patients with moyamoya disease using SELDI-TOF-MS. *BMC Neurol.* **10**:112.
- Bachoud-Lévi A.C., Gaura V., Brugières P., Lefaucheur J.P., Boissé M.F., Maison P., Baudic S., Ribeiro M.J., Bourdet C., Remy P., Cesaro P., Hantraye P., and Peschanski M. (2006) Effect of fetal neural transplants in patients with Huntington's disease 6 years after surgery: a long-term follow-up study. *Lancet. Neurol.* **5**:303-9.
- Barski A., Cuddapah S., Cui K., Roh T.Y., Schones D.E., Wang Z., Wei G., Chepelev I., and Zhao K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell* **129**:823-37.
- Beigneux A.P., Vergnes L., Qiao X., Quatela S., Davis R., Watkins S.M., Coleman R.A., Walzem R.L., Philips M., Reue K., and Young S.G. (2006) Agpat6--a novel lipid biosynthetic gene required for triacylglycerol production in mammary epithelium. *J. Lipid. Res.* **47**:734-44.
- Bell C.G., Finer S., Lindgren C.M., Wilson G.A., Rakyan V.K., Teschendorff A.E., Akan P., Stupka E., Down T.A., Prokopenko I., Morison I.M., Mill J., Pidsley R., International Type 2 Diabetes 1q Consortium, Deloukas P., Frayling T.M., Hattersley A.T., McCarthy M.I., Beck S., and Hitman G.A. (2010) Integrated genetic and epigenetic analysis identifies haplotype-specific methylation in the FTO type 2 diabetes and obesity susceptibility locus. *PLoS One* **5**:e14040.
- Benn C.L., Sun T., Sadri-Vakili G., McFarland K.N., DiRocco D.P., Yohrling G.J., Clark T.W., Bouzou B., and Cha J.H. (2008) Huntingtin modulates transcription, occupies gene promoters in vivo, and binds directly to DNA in a polyglutamine-dependent manner. *J. Neurosci.* **28**:10720-10733.
- Benoit-Marand M., Jaber M., and Gonon F. (2000) Release and elimination of dopamine in vivo in mice lacking the dopamine transporter: functional consequences. *Eur. J. Neurosci.* **12**:2985-92.
- Benson D.A., Karsch-Mizrachi I., Lipman D.J., Ostell J., and Sayers E.W. (2011) GenBank. *Nucleic Acids Res.* **39**:D32-7.

- Berardelli A., Noth J., Thompson P.D., Bollen E.L., Currà A., Deuschl G., van Dijk J.G., Töpper R., Schwarz M., and Roos R.A. (1999) Pathophysiology of chorea and bradykinesia in Huntington's disease. *Mov. Disord.* **14**:398-403.
- Birney E. (2011) Chromatin and heritability: how epigenetic studies can complement genetic approaches. *Trends Genet.* **27**:172-6.
- Bjorklund A., and Stenevi U. (1979) Reconstruction of the nigrostriatal dopamine pathway by intracerebral nigral transplants. *Brain. Res.* **177**:555-560.
- Bock C., Tomazou E.M., Brinkman A.B., Müller F., Simmer F., Gu H., Jäger N., Gnirke A., Stunnenberg H.G., and Meissner A. (2010) Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nat. Biotechnol.* **28**:1106-14.
- Boeneman K., and Crooke E. (2005) Chromosomal replication and the cell membrane. *Curr. Opin. Microbiol.* **8**:143-8.
- Bogdanov M., Mileykovskaya E., and Dowhan W. (2008) Lipids in the assembly of membrane proteins and organization of protein supercomplexes: implications for lipid-linked disorders. *Subcell. Biochem.* **49**:197-239.
- Bogdanov M., Umeda M., and Dowhan W. (1999) Phospholipid-assisted refolding of an integral membrane protein. Minimum structural features for phosphatidylethanolamine to act as a molecular chaperone. *J. Biol. Chem.* **274**:12339-45.
- Bohnen N.I., Kuwabara H., Constantine G.M., Mathis C.A., and Moore R.Y. (2007) Grooved pegboard test as a biomarker of nigrostriatal denervation in Parkinson's disease. *Neurosci. Lett.* **424**:185-9.
- Bonham L., Leung D.W., White T., Hollenback D., Klein P., Tulinsky J., Coon M., de Vries P., and Singer J.W. (2003) Lysophosphatidic acid acyltransferase-beta: a novel target for induction of tumour cell apoptosis. *Expert Opin. Ther. Targets* **7**:643-661.
- Borrell-Pages M., Zala D., Humbert S., and Saudou F. (2006) Huntington's disease: from huntingtin function and dysfunction to therapeutic strategies. *Cell. Mol. Life Sci.* **63**:2642-60.
- Boudreau R.L., McBride J.L., Martins I., Shen S., Xing Y., Carter B.J., and Davidson B.L. (2009) Nonallele-specific silencing of mutant and wild-type huntingtin demonstrates therapeutic efficacy in Huntington's disease mice. *Mol. Ther.* **17**:1053-1063.
- Bouyer J.J., Park D.H., Joh T.H., and Pickel V.M. (1984) Chemical and structural analysis of the relation between cortical inputs and tyrosine hydroxylase-containing terminals in rat neostriatum. *Brain Res.* **302**:267-75.

- Braak H., Del Tredici K., Bratzke H., Hamm-Clement J., Sandmann-Keil D., and Rüb U. (2002) Staging of the intracerebral inclusion body pathology associated with idiopathic Parkinson's disease (preclinical and clinical stages). *J. Neurol.* 249 Suppl **3**:III/1-5.
- Brandt J. Cognitive impairments in Huntington's disease: insights into the neuropsychology of the striatum. In: Boller FaG, J, editor. *handbook of neuropsychology*. Vol 5: Elseviers Sciences Publishers B.V., **1991**: 241-263.
- Brochet X., Lefranc M.P., and Giudicelli V. (2008) IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res.* **36**:W503-8. ([http://www.imgt.org/IMGT\\_vquest/share/textes/](http://www.imgt.org/IMGT_vquest/share/textes/))
- Brochier C., Gaillard M.C., Diguët E., Caudy N., Dossat C., Ségurens B., Wincker P., Roze E., Caboche J., Hantraye P., Brouillet E., Elalouf J.M., and de Chaldée M. (2008) Quantitative gene expression profiling of mouse brain regions reveals differential transcripts conserved in human and affected in disease models. *Physiol. Genomics* **33**:170-9.
- Browne S.E., Bowling A.C., MacGarvey U., Baik M.J., Berger S.C., Muqit M.M., Bird E.D., and Beal M.F. (1997) Oxidative damage and metabolic dysfunction in Huntington's disease: selective vulnerability of the basal ganglia. *Ann. Neurol.* **41**:646-653.
- Bruyn G.W., and Went L.N. Huntington's chorea. In: Vinken P.J, Bruyn G.W., Klawans H.L., editor. *Handbook of clinical Neurology*. Vol 5: Elsevier Science Publisher B.V., **1986**: 267-313.
- Bunney B.S., Walters J.R., Roth R.H., and Aghajanian G.K. (1973) Dopaminergic neurons: effect of antipsychotic drugs and amphetamine on single cell activity. *J. Pharmacol. Exp. Ther.* **185**:560-71.
- Butcher L.M., Meaburn E., Knight J., Sham P.C., Schalkwyk L.C., Craig I.W., and Plomin R. (2005) SNPs, microarrays and pooled DNA: identification of four loci associated with mild mental impairment in a sample of 6000 children. *Hum. Mol. Genet.* **14**:1315-25.
- Caine E.D., and Fisher J.M. Dementia in Huntington's disease. In: Frederiks JAM, editor. *Handbook of clinical neurology*. Vol 2: Elsevier Science Publisher B.V., **1985**: 305-310.
- Cao J., Li J.L., Li D., Tobin J.F., and Gimeno R.E. (2006) Molecular identification of microsomal acyl CoA:glycerol-3-phosphate acyltransferase, a key enzyme in de novo triacylglycerol synthesis. *Proc. Natl. Acad. Sci. USA* **103**:19695-19700.

- Cao J., Liu Y., Lockwood J., Burn P., and Shi Y. (2004) A novel cardiolipin remodeling pathway revealed by a gene encoding an endoplasmic reticulum- associated acyl-CoA:lysocardiolipin acyltransferase (ALCAT1) in mouse. *J. Biol. Chem.* **279**:31727-31734.
- Cao J., Shan D., Revett T., Li D., Wu L., Liu W., Tobin J.F., Gimeno R.E. (2008) Molecular identification of a novel mammalian brain isoform of acyl-CoA:lysophospholipid acyltransferase with prominent ethanolamine lysophospholipid acylating activity, LPEAT2. *J. Biol. Chem.* **283**:19049-57.
- Cao J., Shen W., Chang Z., and Shi Y. (2009) ALCAT1 is a polyglycerophospholipid acyltransferase potently regulated by adenine nucleotide and thyroid status. *Am. J. Physiol. Endocrinol. Metab.* **296**:E647-53.
- Capecchi M.R. (1989) Altering the genome by homologous recombination. *Science* **244**:1288-92.
- Carlson C.M., Dupuy A.J., Fritz S., Roberg-Perez K.J., Fletcher C.F., and Largaespada D.A. (2003) Transposon mutagenesis of the mouse germline. *Genetics* **165**:243-56.
- Carter G.W., Rush C.G., Uygun F., Sakhanenko N.A., Galas D.J., and Galitski T. (2010) A systems-biology approach to modular genetic complexity. *Chaos* **20**:026102.
- Cauli B., Porter J.T., Tsuzuki K., Lambolez B., Rossier J., Quenet B., and Audinat E. (2000) Classification of fusiform neocortical interneurons based on unsupervised clustering. *Proc. Natl. Acad. Sci. USA* **97**:6144-9.
- Charvin D. Dopamine et dégénérescence des neurones striataux dans la maladie de Huntington : vers l'identification de nouvelles cibles thérapeutiques. *Neurosciences*. Paris: Université de Paris VI, **2005**: 216.
- Charvin D., Vanhoutte P., Pages C., Borrelli E., and Caboche J. (2005) Unraveling a role for dopamine in Huntington's disease: the dual role of reactive oxygen species and D2 receptor stimulation. *Proc. Natl. Acad. Sci. USA* **102**: 12218-23.
- Chen J. (2004) Novel regulatory mechanisms of mTOR signaling. *Curr. Top. Microbiol. Immunol.* **279**:245-57.
- Chen X., Hyatt B.A., Mucenski M.L., Mason R.J., and Shannon J.M. (2006) Identification and characterization of a lysophosphatidylcholine acyltransférase in alveolar type II cells. *Proc. Natl. Acad. Sci. USA* **103**:11724-11729.
- Chen Y.Q., Kuo M.S., Li S., Bui H.H., Peake D.A., Sanders P.E., Thibodeaux S.J., Chu S, Qian Y.W., Zhao Y., Bredt D.S., Moller D.E., Konrad R.J., Beigneux A.P., Young

- S.G., and Cao G. (2008) AGPAT6 is a novel microsomal glycerol-3-phosphate acyltransferase. *J. Biol. Chem.* **283**:10048-57.
- Chotkowski H.L., Ciota A.T., Jia Y., Puig-Basagoiti F., Kramer L.D., Shi P.Y., and Glaser R.L. (2008) West Nile virus infection of *Drosophila melanogaster* induces a protective RNAi response. *Virology* **377**:197-206.
- Choudhury B.K., Shi X.Z., and Sarna S.K. (2009) Norepinephrine mediates the transcriptional effects of heterotypic chronic stress on colonic motor function. *Am. J. Physiol. Gastrointest. Liver. Physiol.* **296**:G1238-47.
- Ciccimaro E., and Blair I.A. (2010) Stable-isotope dilution LC–MS for quantitative biomarker analysis. *Bioanalysis* **2**:311-41.
- Ciliax B.J., Drash G.W., Staley J.K., Haber S., Mobley C.J., Miller G.W., Mufson E.J., Mash D.C., and Levey A.I. (1999) Immunocytochemical localization of the dopamine transporter in human brain. *J. Comp. Neurol.* **409**:38-56.
- Cloonan N., Forrest A.R., Kolle G., Gardiner B.B., Faulkner G.J., Brown M.K., Taylor D.F., Steptoe A.L., Wani S., Bethel G., Robertson A.J., Perkins A.C., Bruce S.J., Lee C.C., Ranade S.S., Peckham H.E., Manning J.M., McKernan K.J., and Grimmond S.M. (2008) Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nat. Methods* **5**:613-9.
- Coleman R.A., and Lee D.P. (2004) Enzymes of triacylglycerol synthesis and their regulation. *Prog. Lipid Res.* **43**:134-76.
- Coon M., Ball A., Pound J., Ap S., Hollenback D., White T., Tulinsky J., Bonham L., Morrison D.K., Finney R., and Singer J.W. (2003) Inhibition of lysophosphatidic acid acyltransferase beta disrupts proliferative and survival signals in normal cells and induces apoptosis of tumor cells. *Mol. Cancer Ther.* **2**:1067–1078.
- Cortes V.A., Curtis D.E., Sukumaran S., Shao X., Parameswara V., Rashid S., Smith A.R., Ren J., Esser V., Hammer R.E., Agarwal A.K., Horton J.D., and Garg A. (2009) Molecular mechanisms of hepatic steatosis and insulin resistance in the AGPAT2-deficient mouse model of congenital generalized lipodystrophy. *Cell Metab.* **9**:165-176.
- Dahlström A., and Fuxe K. (1964) Localization of monoamines in the lower brain stem. *Experientia* **20**:398-9.
- Davies S.W., Turmaine M., Cozens B.A., DiFiglia M., Sharp A.H., Ross C.A., Scherzinger E., Wanker E.E., Mangiarini L., and Bates G.P. (1997) Formation of neuronal

- intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation. *Cell* **90**:537-48.
- de Almeida L.P., Ross C.A., Zala D., Aebischer P., and Déglon N. (2002) Lentiviral-mediated delivery of mutant huntingtin in the striatum of rats induces a selective neuropathology modulated by polyglutamine repeat size, huntingtin expression levels, and protein length. *J. Neurosci.* **22**:3473-83.
- de Chaldée M., Brochier C., Van de Vel A., Caudy N., Luthi-Carter R., Gaillard M.C., and Elalouf J.M. (2006) Capucin: a novel striatal marker down-regulated in rodent models of Huntington disease. *Genomics* **87**: 200-7.
- de Chaldée M., Gaillard M.C., Bizat N., Buhler J.M., Manzoni O., Bockaert J., Hantraye P., Brouillet E., and Elalouf J.M. (2003) Quantitative assessment of transcriptome differences between brain territories. *Genome Res.* **13**: 1646-53.
- de la Monte S.M., Vonsattel J.P., and Richardson E.P., Jr. (1988) Morphometric demonstration of atrophic changes in the cerebral cortex, white matter, and neostriatum in Huntington's disease. *J. Neuropathol. Exp. Neurol.* **47**:516-525.
- Deane C.M., Salwiński Ł., Xenarios I., and Eisenberg D. (2002) Protein interactions: two methods for assessment of the reliability of high throughput observations. *Mol. Cell Proteomics.* **1**:349-56.
- Deconinck A.E., Rafael J.A., Skinner J.A., Brown S.C., Potter A.C., Metzinger L., Watt D.J., Dickson J.G., Tinsley J.M., and Davies K.E. (1997) Utrophin-dystrophin-deficient mice as a model for Duchenne muscular dystrophy. *Cell* **90**:717-27.
- deKemp R.A., Ruddy T.D., Hewitt T., Dalipaj M.M., and Beanlands R.S. (2000) Detection of serial changes in absolute myocardial perfusion with <sup>82</sup>Rb PET. *J. Nucl. Med.* **41**:1426-35.
- Delion S., Chalon S., Guilloteau D., Lejeune B., Besnard J.C., and Durand G. (1997) Age-related changes in phospholipid fatty acid composition and monoaminergic neurotransmission in the hippocampus of rats fed a balanced or an n-3 polyunsaturated fatty acid-deficient diet. *J. Lipid. Res.* **38**:680-9.
- DeLong M.R. (1990) Primate models of movement disorders of basal ganglia origin. *Trends Neurosci.* **13**: 281-5.
- Denery J.R., Nunes A.A., Hixon M.S., Dickerson T.J., and Janda K.D. (2010) Metabolomics-based discovery of diagnostic biomarkers for onchocerciasis. *PLoS Negl. Trop. Dis.* **4**:pii:e834.

- Deng W., Wang D.A., Gosmanova E., Johnson L.R., and Tigyi G. (2003) LPA protects intestinal epithelial cells from apoptosis by inhibiting the mitochondrial pathway. *Am. J. Physiol. Gastrointest. Liver. Physiol.* **284**:G821-9.
- Desplats P.A., Kass K.E., Gilmartin T., Stanwood G.D., Woodward E.L., Head S.R., Sutcliffe J.G., and Thomas E.A. (2006) Selective deficits in the expression of striatal-enriched mRNAs in Huntington's disease. *J. Neurochem.* **96**:743-57.
- Di Chiara G., and Imperato A. (1988) Drugs abused by humans preferentially increase synaptic dopamine concentrations in the mesolimbic system of freely moving rats. *Proc. Natl. Acad. Sci. USA* **85**:5274-8.
- Di Maio L., Squitieri F., Napolitano G., Campanella G., Trofatter J.A., and Conneally P.M. (1993) Suicide risk in Huntington's disease. *J. Med. Genet.* **30**:293-5.
- DiFiglia M., Sena-Esteves M., Chase K., Sapp E., Pfister E., Sass M., Yoder J., Reeves P., Pandey R.K., Rajeev K.G., Manoharan M., Sah D.W., Zamore P.D., and Aronin N. (2007) Therapeutic silencing of mutant huntingtin with siRNA attenuates striatal and cortical neuropathology and behavioral deficits. *Proc. Natl. Acad. Sci. USA* **104**:17204-17209.
- Dircks L.K., Ke J., and Sul H.S. (1999) A conserved seven amino acid stretch important for murine mitochondrial glycerol-3-phosphate acyltransferase activity. Significance of arginine 318 in catalysis. *J. Biol. Chem.* **274**:34728-34.
- Divac I., Fonnum F., and Storm-Mathisen J. (1977) High affinity uptake of glutamate in terminals of corticostriatal axons. *Nature* **266**:377-378.
- Donoghue J.P., and Herkenham M. (1986). Neostriatal projections from individual cortical fields conform to histochemically distinct striatal compartments in the rat. *Brain Res.* **365**:397-403.
- Douvas M.G., Hogan K.N., Ji Y., Hollenback D., Bonham L., Singer J.W., and Mitchell B.S. (2006) Effect of lysophosphatidic acid acyltransferase-beta inhibition in acute leukemia. *Leuk. Res.* **30**:1027-1036.
- Doyle J.P., Dougherty J.D., Heiman M., Schmidt E.F., Stevens T.R., Ma G., Bupp S., Shrestha P., Shah R.D., Doughty M.L., Gong S., Greengard P., and Heintz N. (2008) Application of a translational profiling approach for the comparative analysis of CNS cell types. *Cell* **135**:749-62.
- Drouet V., Perrin V., Hassig R., Dufour N., Auregan G., Alves S., Bonvento G., Brouillet E., Luthi-Carter R., Hantraye P., and Déglon N. (2009) Sustained effects of nonallelespecific Huntingtin silencing. *Ann. Neurol.* **65**:276-285.



- Du J., Zhang C., Liu J., Sidky C., and Huang X.P. (2006) A point mutation (R192H) in the C-terminus of human cardiac troponin I causes diastolic dysfunction in transgenic mice. *Arch. Biochem. Biophys.* **456**:143-50.
- Eberhardt C., Gray P.W., and Tjoelker L.W. (1997) Human lysophosphatidic acid acyltransferase. cDNA cloning, expression, and localization to chromosome 9q343. *J. Biol. Chem.* **272**:20299-20305.
- Eberhardt C., Gray P.W., and Tjoelker L.W. (1999) cDNA cloning, expression and chromosomal localization of two human lysophosphatidic acid acyltransferases. *Adv. Exp. Med. Biol.* **469**:351-356.
- English D., Cui Y., and Siddiqui R.A. (1996) Messenger functions of phosphatidic acid. *Chem. Phys. Lipids* **80**:117-132.
- Erickson R.P., McQueen C.A., Chau B., Gokhale V., Uchiyama M., Toyoda A., Ejima F., Maho N., Sakaki Y., and Gondo Y. (2008) An N-ethyl-N-nitrosourea-induced mutation in N-acetyltransferase 1 in mice. *Biochem. Biophys. Res. Commun.* **370**:285-8.
- Faideau M., Kim J., Cormier K., Gilmore R., Welch M., Auregan G., Dufour N., Guillemier M., Brouillet E., Hantraye P., Déglon N., Ferrante R.J., and Bonvento G. (2010) In vivo expression of polyglutamine-expanded huntingtin by mouse striatal astrocytes impairs glutamate transport: a correlation with Huntington's disease subjects. *Hum. Mol. Genet.* **19**:3053-67.
- Feil R., Brocard J., Mascrez B., LeMeur M., Metzger D., and Chambon P. (1996) Ligand-activated site-specific recombination in mice. *Proc. Natl. Acad. Sci. USA* **93**:10887-90.
- Feldman D.J., Frank R.A., Kehne J.H., Flannery R., Brown D., Soni S., Byrd G., and Shah S. (1997) Mixed D2/5-HT2 antagonism differentially affects apomorphine- and amphetamine-induced stereotyped behavior. *Pharmacol. Biochem. Behav.* **58**:565-72.
- Fenn J.B., Mann M., Meng C.K., Wong S.F., and Whitehouse C.M. (1989) Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**:64-71.
- Fields S., and Song O. (1989) A novel genetic system to detect protein-protein interactions. *Nature* **340**:245-6.
- Finn R.D., Mistry J., Tate J., Coghill P., Heger A., Pollington J.E., Gavin O.L., Gunasekaran P., Ceric G., Forslund K., Holm L., Sonnhammer E.L., Eddy S.R., and Bateman A. (2010) The Pfam protein families database. *Nucleic Acids Res.* **38**:D211-22.

- Folstein S.E., Leigh R.J., Parhad I.M., and Folstein M.F. (1986) The diagnosis of Huntington's disease. *Neurology* **36**:1279-1283.
- Ford B. (2010) Pain in Parkinson's disease. *Mov. Disord.* **25**:S98-103.
- Gale S.E., Frolov A., Han X., Bickel P.E., Cao L., Bowcock A., Schaffer J.E., and Ory D.S. (2006) A regulatory role for 1-acylglycerol-3-phosphate-O-acyltransferase 2 in adipocyte differentiation. *J Biol Chem* **281**:11082-11089.
- Garcia M., Vanhoutte P., Pages C., Besson M.J., Brouillet E., and Caboche J. (2002) The mitochondrial toxin 3-nitropropionic acid induces striatal neurodegeneration via a c-Jun N-terminal kinase/c-Jun module. *J. Neurosci.* **22**: 2174-84.
- Ge S., Murugesan N., and Pachter J.S. (2009) Astrocyte- and endothelial-targeted CCL2 conditional knockout mice: critical tools for studying the pathogenesis of neuroinflammation. *J. Mol. Neurosci.* **39**:269-83.
- Gerfen C.R. (1992) The neostriatal mosaic: multiple levels of compartmental organization. *Trends Neurosci.* **15**:133-139.
- Ginsberg L., and Gershfeld N.L. (1991) Membrane bilayer instability and the pathogenesis of disorders of myelin. *Neurosci. Lett.* **130**:133-6.
- Goldman-Rakic P.S. (1995) Cellular basis of working memory. *Neuron* **14**:477-85.
- Gong S., Zheng C., Doughty M.L., Losos K., Didkovsky N., Schambra U.B., Nowak N.J., Joyner A., Leblanc G., Hatten M.E., and Heintz N. (2003) A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* **425**:917-25.
- Gorenstein E.E., Mammato C.A., and Sandy J.M. (1989) Performance of inattentive-overactive children on selected measures of prefrontal-type function. *J. Clin. Psychol.* **45**:619-32.
- Gossler A., Joyner A.L., Rossant J., and Skarnes W.C. (1989) Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. *Science* **244**:463-5.
- Goyenvallé A., Babbs A., Powell D., Kole R., Fletcher S., Wilton S.D., and Davies K.E. (2010) Prevention of dystrophic pathology in severely affected dystrophin/utrophin-deficient mice by morpholino-oligomer-mediated exon-skipping. *Mol. Ther.* **18**:198-205.
- Graybiel A.M., and Ragsdale C.W., Jr. (1978) Histochemically distinct compartments in the striatum of human, monkeys, and cat demonstrated by acetylthiocholinesterase staining. *Proc. Natl. Acad. Sci. USA* **75**:5723-5726.

- Griffiths E.A., and Gore S.D. (2008) DNA methyltransferase and histone deacetylase inhibitors in the treatment of myelodysplastic syndromes. *Semin. Hematol.* **45**:23-30.
- Gu H., Marth J.D., Orban P.C., Mossmann H., and Rajewsky K. (1994) Deletion of a DNA polymerase beta gene segment in T cells using cell type-specific gene targeting. *Science* **265**:103-6.
- Gusella J.F., Wexler N.S., Conneally P.M., Naylor S.L., Anderson M.A., Tanzi R.E., Watkins P.C., Ottina K., Wallace M.R., Sakaguchi A.Y., Young A.B., Shoulson I., Bonilla E., and Martin J.B. (1983) A polymorphic DNA marker genetically linked to Huntington's disease. *Nature* **306**:234-238.
- Haley B., Foys B., and Levine M. (2010) Vectors and parameters that enhance the efficacy of RNAi-mediated gene disruption in transgenic *Drosophila*. *Proc. Natl. Acad. Sci. USA* **107**:11435-40.
- Handley O.J., Naji J.J., Dunnett S.B., and Rosser A.E. (2006) Pharmaceutical, cellular and genetic therapies for Huntington's disease. *Clin. Sci.* **110**:73-88.
- Haque W., Garg A., and Agarwal A.K. (2005) Enzymatic activity of naturally occurring 1-acylglycerol-3-phosphate-O-acyltransferase 2 mutants associated with congenital generalized lipodystrophy. *Biochem. Biophys. Res. Commun.* **327**:446-453.
- Harjes P., and Wanker E.E. (2003) The hunt for huntingtin function: interaction partners tell many different stories. *Trends Biochem. Sci* **28**: 425-33.
- Harper S.Q., Staber P.D., He X., Eliason S.L., Martins I.H., Mao Q., Yang L., Kotin R.M., Paulson H.L., and Davidson B.L. (2005) RNA interference improves motor and neuropathological abnormalities in a Huntington's disease mouse model. *Proc. Natl. Acad. Sci. USA* **102**:5820-5825.
- Harvey B.K., Wang Y., and Hoffer B.J. (2008) Transgenic rodent models of Parkinson's disease. *Acta Neurochir. Suppl.* **101**:89-92.
- Hegele R.A., Joy T.R., Al-Attar S.A., and Rutt B.K. (2007) Thematic review series: Adipocyte Biology. Lipodystrophies: windows on adipose biology and metabolism. *J. Lipid. Res.* **48**:1433-1444.
- Heiman M., Schaefer A., Gong S., Peterson J.D., Day M., Ramsey K.E., Suárez-Fariñas M., Schwarz C., Stephan D.A., Surmeier D.J., Greengard P., and Heintz N. (2008) A translational profiling approach for the molecular characterization of CNS cell types. *Cell* **135**:738-48.

- Heimer L., Zaborszky L., Zahm D.S., and Alheid G.F. (1987) The ventral striatopallidothalamic projection: I. The striatopallidal link originating in the striatal parts of the olfactory tubercle. *J. Comp. Neurol.* **255**:571-91.
- Henneberry A.L., Wright M.M., and McMaster C.R. (2002) The major sites of cellular phospholipid synthesis and molecular determinants of Fatty Acid and lipid head group specificity. *Mol. Biol. Cell.* **13**:3148-61.
- Hersch S.M., and Rosas H.D. (2008) Neuroprotection for Huntington's disease: ready, set, slow. *Neurotherapeutics* **5**: 226-36.
- Hicks S.L., M PAR, C VPG, Tabrizi S.J., and Kennard C. (2008) Oculomotor deficits indicate the progression of Huntington's Disease. *Prog. Brain. Res.* **171**: 555-8.
- Hideshima T., Chauhan D., Ishitsuka K., Yasui H., Raje N., Kumar S., Podar K., Mitsiades C., Hideshima H., Bonham L., Munshi N.C., Richardson P.G., Singer J.W., and Anderson K.C. (2005) Molecular characterization of PS-341 (bortezomib) resistance: implications for overcoming resistance using lysophosphatidic acid acyltransferase (LPAAT)-beta inhibitors. *Oncogene* **24**:3121-3129.
- Hildrestrand G.A., Neurauter C.G., Diep D.B., Castellanos C.G., Krauss S., Bjørås M., and Luna L. (2009) Expression patterns of Neil3 during embryonic brain development and neoplasia. *BMC Neurosci.* **10**:45.
- Hirsch E., Graybiel A.M., and Agid Y.A. (1988) Melanized dopaminergic neurons are differentially susceptible to degeneration in Parkinson's disease. *Nature* **334**:345-8.
- Hishikawa D., Shindou H., Kobayashi S., Nakanishi H., Taguchi R., and Shimizu T. (2008) Discovery of a lysophospholipid acyltransferase family essential for membrane asymmetry and diversity. *Proc. Natl. Acad. Sci. USA* **105**:2830-5.
- Hollenback D., Bonham L., Law L., Rossnagle E., Romero L., Carew H., Tompkins C.K., Leung D.W., Singer J.W., and White T. (2006) Substrate specificity of lysophosphatidic acid acyltransferase beta-evidence from membrane and whole cell assays. *J. Lipid. Res.* **47**:593-604.
- Hoppitt T., Pall H., Calvert M., Gill P., Yao G., Ramsay J., James G., Conduit J., and Sackley C. (2011) A systematic review of the incidence and prevalence of long-term neurological conditions in the UK. *Neuroepidemiology* **36**:19-28.
- Howe P.R., Costa M., Furness J.B., and Chalmers J.P. (1980) Simultaneous demonstration of phenylethanolamine N-methyltransferase immunofluorescent and catecholamine fluorescent nerve cell bodies in the rat medulla oblongata. *Neuroscience* **5**:2229-38.

- Hsieh F.Y., Tengstrand E., Lee J.W., Li L.Y., Silverman L., Riordan B., Miwa G., Milton M., Alden C., and Lee F. (2007) Drug safety evaluation through biomarker analysis--a toxicity study in the cynomolgus monkey using an antibody-cytotoxic conjugate against ovarian cancer. *Toxicol. Appl. Pharmacol.* **224**:12-8.
- Ingvar D.H. (1985) "Memory of the future": an essay on the temporal organization of conscious awareness. *Hum. Neurobiol.* **4**:127-36.
- Ivics Z., Hackett P.B., Plasterk R.H., and Izsvák Z. (1997) Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* **91**:501-10.
- Jaber M., Jones S., Giros B., and Caron M.G. (1997) The dopamine transporter: a crucial component regulating dopamine transmission. *Mov. Disord.* **12**:629-33.
- Jain M., Armstrong R.J., Barker R.A., and Rosser A.E. (2001) Cellular and molecular aspects of striatal development. *Brain Res. Bull.* **55**:533-540.
- Jamsai D., Bianco D.M., Smith S.J., Merriner D.J., Ly-Huynh J.D., Herlihy A., Niranjan B., Gibbs G.M., and O'Bryan M.K. (2008) Characterization of gametogenetin 1 (GGN1) and its potential role in male fertility through the interaction with the ion channel regulator, cysteine-rich secretory protein 2 (CRISP2) in the sperm tail. *Reproduction* **135**:751-9.
- Javoy-Agid F., and Agid Y. (1980) Is the mesocortical dopaminergic system involved in Parkinson disease? *Neurology* **30**:1326-30.
- Jones P.A., and Baylin S.B. (2007) The epigenomics of cancer. *Cell* **128**:683-92.
- Kalashnikova E., Lorca R.A., Kaur I., Barisone G.A., Li B., Ishimaru T., Trimmer J.S., Mohapatra D.P., and Díaz E. (2010) SynDIG1: an activity-regulated, AMPA-receptor-interacting transmembrane protein that regulates excitatory synapse development. *Neuron* **65**:80-93.
- Kalivas P.W., and Duffy P. (1991) A comparison of axonal and somatodendritic dopamine release using in vivo dialysis. *J. Neurochem.* **56**:961-7.
- Kawasaki H., Springett G.M., Toki S., Canales J.J., Harlan P., Blumenstiel J.P., Chen E.J., Bany I.A., Mochizuki N., Ashbacher A., Matsuda M., Housman D.E., and Graybiel A.M. (1998) A Rap guanine nucleotide exchange factor enriched highly in the basal ganglia. *Proc. Natl. Acad. Sci. USA* **95**:13278-83.
- Kegel K.B., Kim M., Sapp E., McIntyre C., Castaño J.G., Aronin N., and DiFiglia M. (2000) Huntingtin expression stimulates endosomal-lysosomal activity, endosome tubulation, and autophagy. *J. Neurosci.* **20**:7268–7278.

- Kegel K.B., Sapp E., Alexander J., Valencia A., Reeves P., Li X., Masso N., Sobin L., Aronin N., and DiFiglia M. (2009) Polyglutamine expansion in huntingtin alters its interaction with phospholipids. *J. Neurochem.* **110**:1585–1597.
- Kellendonk C., Tronche F., Monaghan A.P., Angrand P.O., Stewart F., and Schütz G. (1996) Regulation of Cre recombinase activity by the synthetic steroid RU 486. *Nucleic Acids Res.* **24**:1404-11.
- Kelley L.A., and Sternberg M.J. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* **4**:363-71.
- Kemp J.M., and Powell T.P. (1970) The cortico-striate projection in the monkey. *Brain* **93**: 525-46.
- Kennedy E.P., and Weiss S.B. (1956) The function of cytidine coenzymes in the biosynthesis of phospholipides. *J. Biol. Chem.* **222**:193-214.
- Kent C. (1995) Eukaryotic phospholipid biosynthesis. *Annu. Rev. Biochem.* **64**:315-43.
- Kim D.H., and Rossi J.J. 2007 Strategies for silencing human disease using RNA interference. *Nat. Rev. Genet.* **8**:173-84.
- Kim M., Lee S.T., Chu K., and Kim S.U. (2008) Stem cell-based cell therapy for Huntington disease: a review. *Neuropathology* **28**: 1-9.
- Kirov G., Murphy K.C., Arranz M.J., Jones I., McCandles F., Kunugi H., Murray R.M., McGuffin P., Collier D.A., Owen M.J., and Craddock N. (1998) Low activity allele of catechol-O-methyltransferase gene associated with rapid cycling bipolar disorder. *Mol. Psychiatry* **3**:342-5.
- Kistner A., Gossen M., Zimmermann F., Jerecic J., Ullmer C., Lübbert H., and Bujard H. (1996) Doxycycline-mediated quantitative and tissue-specific control of gene expression in transgenic mice. *Proc. Natl. Acad. Sci. USA* **93**:10933-8.
- Kitada T., Asakawa S., Hattori N., Matsumine H., Yamamura Y., Minoshima S., Yokochi M., Mizuno Y., and Shimizu N. (1998) Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* **392**:605-8.
- Klement I.A., Skinner P.J., Kaytor M.D., Yi H., Hersch S.M., Clark H.B., Zoghbi H.Y., and Orr H.T. (1998) Ataxin-1 nuclear localization and aggregation: role in polyglutamine-induced disease in SCA1 transgenic mice. *Cell* **95**:41-53.
- Koche R.P., Smith Z.D., Adli M., Gu H., Ku M., Gnirke A., Bernstein B.E., and Meissner A. (2011) Reprogramming factor expression initiates widespread targeted chromatin remodeling. *Cell Stem Cell* **8**:96-105.

- Kossowska B., Dudka I., Bugla-Płoskońska G., Szymańska-Chabowska A., Doroszkiewicz W., Gancarz R., Andrzejak R., and Antonowicz-Juchniewicz J. (2010) Proteomic analysis of serum of workers occupationally exposed to arsenic, cadmium, and lead for biomarker research: a preliminary study. *Sci. Total. Environ.* **408**:5317-24.
- Kume K., and Shimizu T. (1997) cDNA cloning and expression of murine 1-acylsn-glycerol-3-phosphate acyltransferase. *Biochem. Biophys. Res. Commun.* **237**:663-666.
- La Rosée P., Jia T., Demehri S., Härtel N., de Vries P., Bonham L., Hollenback D., Singer J.W., Melo J.V., Druker B.J., and Deininger M.W. (2006) Antileukemic activity of lysophosphatidic acid acyltransferase-beta inhibitor CT32228 in chronic myelogenous leukemia sensitive and resistant to imatinib. *Clin. Cancer. Res.* **12**:6540-6546.
- Lands W.E. (1958) Metabolism of glycerolipides; a comparison of lecithin and triglyceride synthesis. *J. Biol. Chem.* **231**:883-8.
- Lein E.S., Hawrylycz M.J., Ao N., Ayres M., Bensinger A., Bernard A., Boe A.F., Boguski M.S., Brockway K.S., Byrnes E.J., Chen L., Chen L., Chen T.M., Chin M.C., Chong J., Crook B.E., Czaplinska A., Dang C.N., Datta S., Dee N.R., Desaki A.L., Desta T., Diep E., Dolbeare T.A., Donelan M.J., Dong H.W., Dougherty J.G., Duncan B.J., Ebbert A.J., Eichele G., Estin L.K., Faber C., Facer B.A., Fields R., Fischer S.R., Fliss T.P., Frensley C., Gates S.N., Glattfelder K.J., Halverson K.R., Hart M.R., Hohmann J.G., Howell M.P., Jeung D.P., Johnson R.A., Karr P.T., Kawal R., Kidney J.M., Knapik R.H., Kuan C.L., Lake J.H., Laramée A.R., Larsen K.D., Lau C., Lemon T.A., Liang A.J., Liu Y., Luong L.T., Michaels J., Morgan J.J., Morgan R.J., Mortrud M.T., Mosqueda N.F., Ng L.L., Ng R., Orta G.J., Overly C.C., Pak T.H., Parry S.E., Pathak S.D., Pearson O.C., Puchalski R.B., Riley Z.L., Rockett H.R., Rowland S.A., Royall J.J., Ruiz M.J., Sarno N.R., Schaffnit K., Shapovalova N.V., Sivisay T., Slaughterbeck C.R., Smith S.C., Smith K.A., Smith B.I., Sodt A.J., Stewart N.N., Stumpf K.R., Sunkin S.M., Sutram M., Tam A., Teemer C.D., Thaller C., Thompson C.L., Varnam L.R., Visel A., Whitlock R.M., Wohnoutka P.E., Wolkey C.K., Wong V.Y., Wood M., Yaylaoglu M.B., Young R.C., Youngstrom B.L., Yuan X.F., Zhang B., Zwingman T.A., Jones A.R. (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**:168-76.
- Letarte M., Voulgaraki D., Hatherley D., Foster-Cuevas M., Saunders N.J., and Barclay A.N. (2005) Analysis of leukocyte membrane protein interactions using protein microarrays. *BMC Biochem.* **6**:2.

- Leung D.W. (2001) The structure and functions of human lysophosphatidic acid acyltransferases. *Front. Biosci.* **6**:D944-D953.
- Lewin T.M., Wang P., and Coleman R.A. (1999) Analysis of amino acid motifs diagnostic for the *sn*-glycerol-3-phosphate acyltransferase reaction. *Biochemistry* **38**:5764-71.
- Li D., Yu L., Wu H., Shan Y., Guo J., Dang Y., Wei Y., and Zhao S. (2003) Cloning and identification of the human LPAAT-zeta gene, a novel member of the lysophosphatidic acid acyltransferase family. *J. Hum. Genet.* **48**:438-42.
- Li J., Romestaing C., Han X., Li Y., Hao X., Wu Y., Sun C., Liu X., Jefferson L.S., Xiong J., Lanoue K.F., Chang Z., Lynch C.J., Wang H., and Shi Y. (2010) Cardiolipin remodeling by ALCAT1 links oxidative stress and mitochondrial dysfunction to obesity. *Cell Metab.* **12**:154-65.
- Liang S., Li Y., Be X., Howes S., and Liu W. (2006) Detecting and profiling tissue-selective genes. *Physiol. Genomics* **26**:158-62.
- Lindbo J.A., Silva-Rosales L., Proebsting, W.M., and Dougherty W.G. (1993) Induction of a highly specific antiviral state in transgenic plants--implications for regulation of gene expression and virus resistance. *Plant Cell* **5**:1749-1759
- Lingford-Hughes A., and Nutt D. (2003) Neurobiology of addiction and implications for treatment. *Br. J. Psychiatry* **182**:97-100.
- Liu B., Xu H., Paton J.F., and Kasparov S. (2010) Cell- and region-specific miR30-based gene knock-down with temporal control in the rat brain. *BMC Mol. Biol.* **11**:93.
- Liu P., Jenkins N.A., and Copeland N.G. (2003) A highly efficient recombineering-based method for generating conditional knockout mutations. *Genome Res.* **13**:476-84.
- Liu Y.F. (1998) Expression of polyglutamine-expanded Huntingtin activates the SEK1-JNK pathway and induces apoptosis in a hippocampal neuronal cell line. *J. Biol. Chem.* **273**: 28873-7.
- Lombardi M.S., Jaspers L., Spronkmans C., Gellera C., Taroni F., Di Maria E., Donato S.D., and Kaemmerer W.F. (2009) A majority of Huntington's disease patients may be treatable by individualized allele-specific RNA interference. *Exp. Neurol.* **217**:312-319.
- Lovinger D.M., White G., and Weight F.F. (1989) Ethanol inhibits NMDA-activated ion current in hippocampal neurons. *Science* **243**:1721-1724
- Lu B., Jiang Y.J., Man M.Q., Brown B., Elias P.M., and Feingold K.R. (2005) Expression and regulation of 1-acyl-*sn*-glycerol- 3-phosphate acyltransferases in the epidermis. *J. Lipid. Res.* **46**:2448-57.



- Lu B., Jiang Y.J., Zhou Y., Xu F.Y., Hatch G.M., and Choy P.C. (2005) Cloning and characterization of murine 1-acyl-*sn*-glycerol 3-phosphate acyltransferases and their regulation by PPAR $\alpha$  in murine heart. *Biochem. J.* **385**:469-477.
- Machida Y., Okada T., Kurosawa M., Oyama F., Ozawa K., and Nukina N. (2006) rAAV-mediated shRNA ameliorated neuropathology in Huntington disease model mouse. *Biochem. Biophys. Res. Commun.* **343**:190-197.
- Magdaleno S., Jensen P., Brumwell C.L., Seal A., Lehman K., Asbury A., Cheung T., Cornelius T., Batten D.M., Eden C., Norland S.M., Rice D.S., Dossooye N., Shakya S., Mehta P., and Curran T. (2006) BGEM: an in situ hybridization database of gene expression in the embryonic and adult mouse nervous system. *PLoS Biol.* **4**:e86.
- Mahant N., McCusker E.A., Byth K., and Graham S. (2003) Huntington's disease: clinical correlates of disability and progression. *Neurology* **61**:1085-1092.
- Maharajan P., Maharajan V., Ravagnan G., and Paino G. (2001) The weaver mutant mouse: a model to study the ontogeny of dopamine transmission systems and their role in drug addiction. *Prog. Neurobiol.* **64**:269-76.
- Makarov A., Denisov E., Kholomeev A., Balschun W., Lange O., Strupat K., and Horning S. (2006) Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Anal. Chem.* **78**:2113-20.
- Makide K., Kitamura H., Sato Y., Okutani M., and Aoki J. (2009) Emerging lysophospholipid mediators, lysophosphatidylserine, lysophosphatidylthreonine, lysophosphatidylethanolamine and lysophosphatidylglycerol. *Prostaglandins Other Lipid Mediat.* **89**:135-9.
- Mangiarini L., Sathasivam K., Seller M., Cozens B., Harper A., Hetherington C., Lawton M., Trotter Y., Lehrach H., Davies S.W., and Bates G.P. (1996) Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell* **87**:493-506.
- Manolio T.A. (2010) Genomewide association studies and assessment of the risk of disease. *N. Engl. J. Med.* **363**:166-76.
- Mansilla, F., da Costa K.A., Wang S., Kruhoffer M., Lewin T.M., Orntoft T.F., Coleman R.A., and Birkenkamp-Demtroder K. (2009) Lysophosphatidylcholine acyltransferase 1 (LPCAT1) overexpression in human colorectal cancer. *J. Mol. Med.* **87**: 85-97.
- Margolis R.L., and Ross C.A. (2001) Expansion explosion: new clues to the pathogenesis of repeat expansion neurodegenerative diseases. *Trends Mol. Med.* **7**: 479-82.

- Marín F., Herrero M.T., Vyas S., and Puellas L. (2005) Ontogeny of tyrosine hydroxylase mRNA expression in mid- and forebrain: neuromeric pattern and novel positive regions. *Dev. Dyn.* **234**:709-17.
- Marks J.D., Hoogenboom H.R., Bonnert T.P., McCafferty J., Griffiths A.D., and Winter G. (1991) By-passing immunization. Human antibodies from V-gene libraries displayed on phage. *J. Mol. Biol.* **222**:581-97.
- Marks J.D., Hoogenboom H.R., Griffiths A.D., and Winter G. (1992) Molecular evolution of proteins on filamentous phage. Mimicking the strategy of the immune system. *J. Biol. Chem.* **267**:16007-10.
- Marsh D. (2007) Lateral pressure profile, spontaneous curvature frustration, and the incorporation and conformation of proteins in membranes. *Biophys. J.* **93**:3884-99
- Martin G.R. (1981) Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. USA* **78**:7634-8.
- Massouh M. Les interneurones geants exprimant la calretinine dans le striatum humain : leur devenir dans la maladie de Huntington. *Neurobiologie*. Vol M.Sc. Québec: Université Laval, **2007**: 76.
- Masuda K., Marasa B., Martindale J.L., Halushka M.K., and Gorospe M. (2009) Tissue- and age-dependent expression of RNA-binding proteins that influence mRNA turnover and translation. *Aging (Albany NY)* **1**:681-98.
- Masuya H., Makita Y., Kobayashi N., Nishikata K., Yoshida Y., Mochizuki Y., Doi K., Takatsuki T., Waki K., Tanaka N., Ishii M., Matsushima A., Takahashi S., Hijikata A., Kozaki K., Furuichi T., Kawaji H., Wakana S., Nakamura Y., Yoshiki A., Murata T., Fukami-Kobayashi K., Mohan S., Ohara O., Hayashizaki Y., Mizoguchi R., Obata Y., and Toyoda T. (2011) The RIKEN integrated database of mammals. *Nucleic Acids Res.* **39**:D861-70.
- Masuya H., Nakai Y., Motegi H., Niinaya N., Kida Y., Kaneko Y., Aritake H., Suzuki N., Ishii J., Koorikawa K., Suzuki T., Inoue M., Kobayashi K., Toki H., Wada Y., Kaneda H., Ishijima J., Takahashi K.R., Minowa O., Noda T., Wakana S., Gondo Y., and Shiroishi T. (2004) Development and implementation of a database system to manage a large-scale mouse ENU-mutagenesis program. *Mamm. Genome* **15**:404-11.
- McBride JL, Boudreau R.L., Harper S.Q., Staber P.D., Monteys A.M., Martins I., Gilmore B.L., Burstein H., Peluso R.W., Polisky B., Carter B.J., and Davidson B.L. (2008)

- Artificial miRNAs mitigate shRNA-mediated toxicity in the brain: implications for the therapeutic development of RNAi. *Proc. Natl. Acad. Sci. USA* **105**:5868-5873.
- Menalled L.B., Sison J.D., Dragatsis I., Zeitlin S., and Chesselet M.F. (2003) Time course of early motor and neuropathological anomalies in a knock-in mouse model of Huntington's disease with 140 CAG repeats. *J. Comp. Neurol.* **465**:11-26.
- Menalled L.B., Sison J.D., Wu Y., Olivieri M., Li X.J., Li H., Zeitlin S., and Chesselet M.F. (2002) Early motor dysfunction and striosomal distribution of huntingtin microaggregates in Huntington's disease knock-in mice. *J. Neurosci.* **22**:8266-76.
- Metzker M.L. (2010) Sequencing technologies - the next generation. *Nat. Rev. Genet.* **11**:31-46.
- Miller G., Neilan M., Chia R., Gheryani N., Holt N., Charbit A., Wells S., Tucci V., Lalanne Z., Denny P., Fisher E.M., Cheeseman M., Askew G.N., and Dear T.N. (2010) ENU mutagenesis reveals a novel phenotype of reduced limb strength in mice lacking fibrillin 2. *PLoS. One.* **5**:e9137.
- Morgan H.D., Sutherland H.G., Martin D.I., and Whitelaw E. (1999) Epigenetic inheritance at the agouti locus in the mouse. *Nat. Genet.* **23**:314-8.
- Morris G., Nevet A., and Bergman H. (2003) Anatomical funneling, sparse connectivity and redundancy reduction in the neural networks of the basal ganglia. *J. Physiol. Paris* **97**: 581-9.
- Mouse Genome Sequencing Consortium, et al. (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**:520-62.
- Nagatsu T. (1989) The human tyrosine hydroxylase gene. *Cell. Mol. Neurobiol.* **9**:313-21.
- Nagle C.A., Vergnes L., Dejong H., Wang S., Lewin T.M., Reue K., and Coleman R.A. (2008) Identification of a novel sn-glycerol-3-phosphate acyltransferase isoform, GPAT4, as the enzyme deficient in *Agpat6*<sup>-/-</sup> mice. *J. Lipid Res.* **49**:823-831.
- Nakanishi H., Shindou H., Hishikawa D., Harayama T., Ogasawara R., Suwabe A., Taguchi R., and Shimizu T. (2006) Cloning and characterization of mouse lung-type acyl-CoA:lysophosphatidylcholine acyltransferase 1 (LPCAT1). Expression in alveolar type II cells and possible involvement in surfactant production. *J. Biol. Chem.* **281**:20140-20147.
- Naldini L., Blömer U., Gallay P., Ory D., Mulligan R., Gage F.H., Verma I.M., and Trono D. (1996) In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**:263-7.

- Nayak S, Herzog RW. 2010 Progress and prospects: immune responses to viral vectors. *Gene Ther.* **17**:295-304.
- Nielsen U.B., Kirpotin D.B., Pickering E.M., Drummond D.C., and Marks J.D. (2006) A novel assay for monitoring internalization of nanocarrier coupled antibodies. *BMC Immunol.* **7**:24.
- Nishi A., Snyder G.L., and Greengard P. (1997) Bidirectional regulation of DARPP-32 phosphorylation by dopamine. *J. Neurosci.* **17**:8147-55.
- Nitsch R.M., Blusztajn J.K., Pittas A.G., Slack B.E., Growdon J.H., and Wurtman R.J. (1992) Evidence for a membrane defect in Alzheimer disease brain. *Proc. Natl. Acad. Sci. USA* **89**:1671-5.
- Ouimet C.C., Miller P.E., Hemmings H.C. Jr, Walaas S.I., and Greengard P. (1984) DARPP-32, a dopamine- and adenosine 3':5'-monophosphate-regulated phosphoprotein enriched in dopamine-innervated brain regions. III. Immunocytochemical localization. *J. Neurosci.* **4**:111-24.
- Paddison P.J., Caudy A.A., Bernstein E., Hannon G.J., and Conklin D.S. (2002) Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes Dev.* **16**:948-58.
- Pagel J.M., Laugen C., Bonham L., Hackman R.C., Hockenbery D.M., Bhatt R., Hollenback D., Carew H., Singer J.W., and Press O.W. (2005) Induction of apoptosis using inhibitors of lysophosphatidic acid acyltransferase-beta and anti-CD20 monoclonal antibodies for treatment of human non- Hodgkin's lymphomas. *Clin. Cancer Res.* **11**:4857-4866.
- Palfi S., Ferrante R.J., Brouillet E., Beal M.F., Dolan R., Guyot M.C., Peschanski M., and Hantraye P. (1996) Chronic 3-nitropropionic acid treatment in baboons replicates the cognitive and motor deficits of Huntington's disease. *J. Neurosci.* **16**:3019-25.
- Panegyres P.K. (2004) The contribution of the study of neurodegenerative disorders to the understanding of human memory. *Qjm.* **97**: 555-67.
- Parent A. (1990) Extrinsic connections of the basal ganglia. *Trends Neurosci.* **13**: 254-8.
- Parent A., and Hazrati L.N. (1995) Functional anatomy of the basal ganglia. I. The cortico-basal ganglia-thalamo-cortical loop. *Brain. Res. Rev.* **20**: 91-127.
- Parent M. La collateralisation axonale dans les ganglions de la base chez le primate. neurobiologie. Vol Ph.D. Québec: Université Laval, **2006**: 312.
- Parkinson J. (1817) *An Essay on the Shaking Palsy.*

- Perlow M.J., Freed W.J., Hoffer B.J., Seiger A., Olson L., and Wyatt R.J. (1979) Brain grafts reduce motor abnormalities produced by destruction of nigrostriatal dopamine system. *Science* **204**:643-647.
- Pfister E.L., Kennington L., Straubhaar J., Wagh S., Liu W., DiFiglia M., Landwehrmeyer B., Vonsattel J.P., Zamore P.D., and Aronin N. (2009) Five siRNAs targeting three SNPs may provide therapy for three-quarters of Huntington's disease patients. *Curr. Biol.* **19**:774-778.
- Pidoplichko V.I., DeBiasi M., Williams J.T., and Dani J.A. (1997) Nicotine activates and desensitizes midbrain dopamine neurons. *Nature* **390**:401-4.
- Polymeropoulos M.H., Lavedan C., Leroy E., Ide S.E., Dehejia A., Dutra A., Pike B., Root H., Rubenstein J., Boyer R., Stenroos E.S., Chandrasekharappa S., Athanassiadou A., Papapetropoulos T., Johnson W.G., Lazzarini A.M., Duvoisin R.C., Di Iorio G., Golbe L.I., and Nussbaum R.L. (1997) Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* **276**:2045-7.
- Potter N.T., Spector E.B., and Prior T.W. (2004) Technical standards and guidelines for Huntington disease testing. *Genet. Med.* **6**: 61-5.
- Prasad S.S., Garg A., and Agarwal A.K. (2011) Enzymatic activities of the human AGPAT isoform 3 and isoform 5: localization of AGPAT5 to mitochondria. *J. Lipid Res.* **52**:451-62.
- Preti A. (2007) New developments in the pharmacotherapy of cocaine abuse. *Addict. Biol.* **12**:133-51.
- Putney J.W. Jr., Weiss S.J., Van De Walle C.M., and Haddas R.A. (1980) Is phosphatidic acid a calcium ionophore under neurohumoral control? *Nature* **284**:345-7.
- Qin J.Y., Zhang L., Clift K.L., Huler I., Xiang A.P., Ren B.Z., and Lahn B.T. (2010) Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *PLoS One* **5**:e10611.
- Quinn B.A., Brake T., Hua X., Baxter-Jones K., Litwin S., Ellenson L.H., and Connolly D.C. (2009) Induction of ovarian leiomyosarcomas in mice by conditional inactivation of Brcal and p53. *PLoS One* **4**:e8404.
- Rangone H., Humbert S., and Saudou F. (2004) Huntington's disease: how does huntingtin, an anti-apoptotic protein, become toxic? *Pathol. Biol.* **52**: 338-42.
- Reddy P.H., Williams M., Charles V., Garrett L., Pike-Buchanan L., Whetsell W.O. Jr, Miller G., and Tagle D.A. (1998) Behavioural abnormalities and selective neuronal loss in

- HD transgenic mice expressing mutated full-length HD cDNA. *Nat. Genet.* **20**:198-202.
- Riekkinen P., Rinne U.K., Pelliniemi T.T., and Sonninen V. (1975) Interaction between dopamine and phospholipids. Studies of the substantia nigra in Parkinson disease patients. *Arch. Neurol.* **32**:25-7.
- Rite I., Argüelles S., Venero J.L., García-Rodríguez S., Ayala A., Cano J., and Machado A. (2007) Proteomic identification of biomarkers in the cerebrospinal fluid in a rat model of nigrostriatal dopaminergic degeneration. *J. Neurosci. Res.* **85**:3607-18.
- Ritz M.C., Lamb R.J., Goldberg S.R., and Kuhar M.J. (1987) Cocaine receptors on dopamine transporters are related to self-administration of cocaine. *Science* **237**:1219-23.
- Rizzo M.A., and Shome K., (2000) Watkins S.C., and Romero G. The recruitment of Raf-1 to membranes is mediated by direct interaction with phosphatidic acid and is independent of association with Ras. *J. Biol. Chem.* **275**:23911-8.
- Robins Wahlin T.B., Bäckman L., Lundin A., Haegermark A., Winblad B., and Anvret M. (2000) High suicidal ideation in persons testing for Huntington's disease. *Acta Neurol. Scand.* **102**:150-61.
- Rodriguez-Lebron E., Denovan-Wright E.M., Nash K., Lewin A.S., and Mandel R.J. (2005) Intrastriatal rAAV-mediated delivery of anti-huntingtin shRNAs induces partial reversal of disease progression in R6/1 Huntington's disease transgenic mice. *Mol. Ther.* **12**:618-633.
- Rosas H.D., Koroshetz W.J., Chen Y.I., Skeuse C., Vangel M., Cudkowicz M.E., Caplan K., Marek K., Seidman L.J., Makris N., Jenkins B.G., and Goldstein J.M. (2003) Evidence for more widespread cerebral pathology in early HD: an MRI-based morphometric analysis. *Neurology* **60**:1615-20.
- Rosen G.D., Chesler E.J., Manly K.F., and Williams R.W. (2007) An informatics approach to systems neurogenetics. *Methods Mol. Biol.* **401**:287-303.
- Rossner M.J., Hirrlinger J., Wichert S.P., Boehm C., Newrzella D., Hiemisch H., Eisenhardt G., Stuenkel C., von Ahsen O., and Nave K.A. (2006) Global transcriptome analysis of genetically identified neurons in the adult cortex. *J. Neurosci.* **26**:9956-66.
- Ruan H., and Pownall H.J. (2001) Overexpression of 1-acyl-glycerol-3-phosphate acyltransferase-alpha enhances lipid storage in cellular models of adipose tissue and skeletal muscle. *Diabetes* **50**:233-240.

- Russell W.L., Hunsicker P.R., Raymer G.D., Steele M.H., Stelzner K.F., and Thompson H.M. (1982) Dose--response curve for ethylnitrosourea-induced specific-locus mutations in mouse spermatogonia. *Proc. Natl. Acad. Sci. USA* **79**:3589-91.
- Saghatelian A., Trauger S.A., Want E.J., Hawkins E.G., Siuzdak G., and Cravatt B.F. (2004) Assignment of endogenous substrates to enzymes by global metabolite profiling. *Biochemistry* **43**:14332-9.
- Saha S., Sparks A.B., Rago C., Akmaev V., Wang C.J., Vogelstein B., Kinzler K.W., and Velculescu V.E. (2002) Using the transcriptome to annotate the genome. *Nat. Biotechnol.* **20**:508-12.
- Salmon D.M., and Honeyman T.W. (1980) Proposed mechanism of cholinergic action in smooth muscle. *Nature* **284**:344-5.
- Sasaki M., Shibata E., Tohyama K., Kudo K., Endoh J., Otsuka K., and Sakai A. (2008) Monoamine neurons in the human brain stem: anatomy, magnetic resonance imaging findings, and clinical implications. *Neuroreport* **19**:1649-54.
- Saudou F., Finkbeiner S., Devys D., and Greenberg M.E. (1998) Huntingtin acts in the nucleus to induce apoptosis but death does not correlate with the formation of intranuclear inclusions. *Cell* **95**:55-66.
- Scatton B., Rouquier L., Javoy-Agid F., and Agid Y. (1982) Dopamine deficiency in the cerebral cortex in Parkinson disease. *Neurology* **32**:1039-40.
- Schadt E.E., Molony C., Chudin E., Hao K., Yang X., Lum P.Y., Kasarskis A., Zhang B., Wang S., Suver C., Zhu J., Millstein J., Sieberts S., Lamb J., GuhaThakurta D., Derry J., Storey J.D., Avila-Campillo I., Kruger M.J., Johnson J.M., Rohl C.A., van Nas A., Mehrabian M., Drake T.A., Lusis A.J., Smith R.C., Guengerich F.P., Strom S.C., Schuetz E., Rushmore T.H., and Ulrich R. (2008) Mapping the genetic architecture of gene expression in human liver. *PLoS Biol.* **6**:e107.
- Schier R., Marks J.D., Wolf E.J., Apell G., Wong C., McCartney J.E., Bookman M.A., Huston J.S., Houston L.L., Weiner L.M., and Adams G.P. (1995) In vitro and in vivo characterization of a human anti-c-erbB-2 single-chain Fv isolated from a filamentous phage antibody library. *Immunotechnology* **1**:73-81.
- Schmidt J.A., and Brown W.J. (2009) Lysophosphatidic acid acyltransferase 3 regulates Golgi complex structure and function. *J. Cell. Biol.* **186**:211-8.
- Schmidt J.A., Yvone G.M., and Brown W.J. (2010) Membrane topology of human AGPAT3 (LPAAT3). *Biochem. Biophys. Res. Commun.* **397**:661-7.

- Schneider J., Lothar A., Hein L., and Gilsbach R. (2011) Chronic cardiac pressure overload induces adrenal medulla hypertrophy and increased catecholamine synthesis. *Basic. Res. Cardiol.* **106**:591-602.
- Schultz J., Copley R.R., Doerks T., Ponting C.P., and Bork P. (2000) SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res.* **28**:231-4.
- Shah S.J., Yu K.H., Sangar V., Parry S.I., and Blair I.A. (2009) Identification and quantification of preterm birth biomarkers in human cervicovaginal fluid by liquid chromatography/tandem mass spectrometry. *J. Proteome Res.* **8**:2407-17.
- Shi R.Q., Lee J.K., Hayashi Y., Takeuchi Y., Kambe F., Futaki S., Seo H., Murata Y., and Kodama I. (2008) Long-term amiodarone treatment causes cardioselective hypothyroid-like alteration in gene expression profile. *Eur. J. Pharmacol.* **578**:270-8.
- Shin J.Y., Fang Z.H., Yu Z.X., Wang C.E., Li S.H., and Li X.J. (2005) Expression of mutant huntingtin in glial cells contributes to neuronal excitotoxicity. *J. Cell. Biol.* **171**: 1001-12.
- Shindou H., and Shimizu T. (2009) Acyl-CoA:lysophospholipid acyltransferases. *J. Biol. Chem.* **284**:1-5.
- Shiwach R.S., and Norbury C.G. (1994) A controlled psychiatric study of individuals at risk for Huntington's disease. *Br. J. Psychiatry* **165**:500-5.
- Simha V., and Garg A. (2003) Phenotypic heterogeneity in body fat distribution in patients with congenital generalized lipodystrophy caused by mutations in the AGPAT2 or seipin genes. *J. Clin. Endocrinol. Metab.* **88**:5433-5437.
- Smith C.A., Want E.J., O'Maille G., Abagyan R., and Siuzdak G. (2006) XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.* **78**:779-87.
- Snel B., Lehmann G., Bork P., and Huynen M.A. (2000) STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic Acids Res.* **28**:3442-4.
- Snowden J.S., Gibbons Z.C., Blackshaw A., Doubleday E., Thompson J., Craufurd D., Foster J., Happé F., and Neary D. (2003) Social cognition in frontotemporal dementia and Huntington's disease. *Neuropsychologia* **41**: 688-701.
- Spillantini M.G., Crowther R.A., Jakes R., Hasegawa M., and Goedert M. (1998) alpha-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with lewy bodies. *Proc. Natl. Acad. Sci. USA* **95**:6469-73.



- Springett G.M., Bonham L., Hummer A., Linkov I., Misra D., Ma C., Pezzoni G., Di Giovine S., Singer J., Kawasaki H., Spriggs D., Soslow R., and Dupont J. (2005) Lysophosphatidic acid acyltransferase-beta is a prognostic marker and therapeutic target in gynaecologic malignancies. *Cancer Res.* **65**:9415-9425.
- Squitieri F., Gellera C., Cannella M., Mariotti C., Cislighi G., Rubinstein D.C., Almqvist E.W., Turner D., Bachoud-Lévi A.C., Simpson S.A., Delatycki M., Maglione V., Hayden M.R., and Donato S.D. (2003) Homozygosity for CAG mutation in Huntington disease is associated with a more severe clinical course. *Brain* **126**:946-55.
- Stamps A.C., Elmore M.A., Hill M.E., Kelly K., Makda A.A., and Finnen M.J. (1997) A human cDNA sequence with homology to non-mammalian lysophosphatidic acid acyltransferases. *Biochem. J.* **326**:455-461.
- Stanford K.I., Wang L., Castagnola J., Song D., Bishop J.R., Brown J.R., Lawrence R., Bai X., Habuchi H., Tanaka M., Cardoso W.V., Kimata K., and Esko J.D. (2010) Heparan sulfate 2-O-sulfotransferase is required for triglyceride-rich lipoprotein clearance. *J. Biol. Chem.* **285**:286-94.
- Stark C., Breitkreutz B.J., Reguly T., Boucher L., Breitkreutz A., and Tyers M. (2006) BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* **34**:D535-9.
- Steffan J.S., Kazantsev A., Spasic-Boskovic O., Greenwald M., Zhu Y.Z., Gohler H., Wanker E.E., Bates G.P., Housman D.E., and Thompson L.M. (2000) The Huntington's disease protein interacts with p53 and CREB-binding protein and represses transcription. *Proc. Natl. Acad. Sci. USA* **97**: 6763-8.
- Strand A.D., Aragaki A.K., Baquet Z.C., Hodges A., Cunningham P., Holmans P., Jones K.R., Jones L., Kooperberg C., and Olson J.M. (2007) Conservation of regional gene expression in mouse and human brain. *PLoS Genet.* **3**:e59.
- Sugars K.L., and Rubinstein D.C. (2003) Transcriptional abnormalities in Huntington disease. *Trends Genet.* **19**: 233-8.
- Sukumaran S., Barnes R., Garg A., and Agarwal A. (2009) Functional characterization of the human 1-acylglycerol-3-phosphate-O-acyltransferase isoform 10/glycerol-3-phosphate acyltransferase isoform 3. *J. Mol. Endocrinol.* **42**:469-478.
- Tabet J.C., and Rebuffat S. (2003) Nobel Prize 2002 for chemistry: mass spectrometry and nuclear magnetic resonance. *Med. Sci.* **19**:865-72.
- Tabibiazar R., Wagner R.A., Deng A., Tsao P.S., and Quertermous T. (2006) Proteomic profiles of serum inflammatory markers accurately predict atherosclerosis in mice. *Physiol. Genomics.* **25**:194-202.

- Talio M.C., Luconi M.O., Masi A.N., and Fernández L.P. (2010) Cadmium monitoring in saliva and urine as indicator of smoking addiction. *Sci. Total. Environ.* **408**:3125-32.
- Tang W., Yuan J., Chen X., Gu X., Luo K., Li J., Wan B., Wang Y., and Yu L. (2006) Identification of a novel human lysophosphatidic acid acyltransferase, LPAAT-theta, which activates mTOR pathway. *J. Biochem. Mol. Biol.* **39**:626-635.
- The American College of Medical Genetics/American Society of Human Genetics Huntington Disease Genetic Testing Working Group. (1998) ACMG/ASHG statement. Laboratory guidelines for Huntington disease genetic testing. *Am. J. Hum. Genet.* **62**:1243-7.
- The Huntington's Disease Collaborative Research Group. (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **72**:971-83.
- The International HapMap consortium. (2005) A haplotype map of the human genome. *Nature* **437**:1299–1320.
- Timchenko L.T., and Caskey C.T. (1996) Trinucleotide repeat disorders in humans: discussions of mechanisms and medical issues. *FASEB J.* **10**:1589-97.
- Tinoco M.L., Dias B.B., Dall'Asta R.C., Pamphile J.A., and Aragão F.J. (2010) In vivo trans-specific gene silencing in fungal cells by in planta expression of a double-stranded RNA. *BMC Biol.* **8**:27.
- Tseveleki V., Rubio R., Vamvakas S.S., White J., Taoufik E., Petit E., Quackenbush J., and Probert L. (2010) Comparative gene expression analysis in mouse models for multiple sclerosis, Alzheimer's disease and stroke for identifying commonly regulated and disease-specific gene changes. *Genomics* **96**:82-91.
- Turner T.T., Bomgardner D., Jacobs J.P., and Nguyen Q.A. (2003) Association of segmentation of the epididymal interstitium with segmented tubule function in rats and mice. *Reproduction* **125**:871-8.
- Uhlén M., Björling E., Agaton C., Szigartyo C.A., Amini B., Andersen E., Andersson A.C., Angelidou P., Asplund A., Asplund C., Berglund L., Bergström K., Brumer H., Cerjan D., Ekström M., Elobeid A., Eriksson C., Fagerberg L., Falk R., Fall J., Forsberg M., Björklund M.G., Gumbel K., Halimi A., Hallin I., Hamsten C., Hansson M., Hedhammar M., Hercules G., Kampf C., Larsson K., Lindskog M., Lodewyckx W., Lund J., Lundeberg J., Magnusson K., Malm E., Nilsson P., Odling J., Oksvold P., Olsson I., Oster E., Ottosson J., Paavilainen L., Persson A., Rimini R., Rockberg J., Runeson M., Sivertsson A., Sköllerö A., Steen J., Stenvall M., Sterky F., Strömberg S., Sundberg M., Tegel H., Tourle S., Wahlund E., Waldén A., Wan J., Wernérus H.,

- Westberg J., Wester K., Wrethagen U., Xu L.L., Hober S., and Pontén F. (2005) A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol. Cell Proteomics* **4**:1920-32.
- UniProt Consortium. (2011) Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Res.* **39**:D214-9.
- Urnov F.D., Rebar E.J., Holmes M.C., Zhang H.S., and Gregory P.D. (2010) Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **11**:636-46.
- van Bilsen P.H., Jaspers L., Lombardi M.S., Odekerken J.C., Burright E.N., and Kaemmerer W.F. (2008) Identification and allele-specific silencing of the mutant huntingtin allele in Huntington's disease patient-derived fibroblasts. *Hum. Gene Ther.* **19**:710-719.
- van der Kooy D. (1979) The organization of the thalamic, nigral and raphe cells projecting to the medial vs lateral caudate-putamen in rat. A fluorescent retrograde double labeling study. *Brain Res.* **169**:381-387.
- van Meer G. (2005) Cellular lipidomics. *EMBO J.* **24**:3159-65.
- van Os J., and Kapur S. (2009) Schizophrenia. *Lancet* **374**:635-45.
- Velculescu V.E., Zhang L., Vogelstein B., and Kinzler K.W. (1995) Serial analysis of gene expression. *Science* **270**: 484-7.
- Vengeliene V., Bilbao A., Molander A., and Spanagel R. (2008) Neuropharmacology of alcohol addiction. *Br. J. Pharmacol.* **154**:299-315.
- Vergnes L., Beigneux A.P., Davis R., Watkins S.M., Young S.G., and Reue K. (2006) Agpat6 deficiency causes subdermal lipodystrophy and resistance to obesity. *Journal of Lipid Research* **47**:745-754.
- Virlon B., Cheval L., Buhler J.M., Billon E., Doucet A., and Elalouf J.M. (1999) Serial microanalysis of renal transcriptomes. *Proc Natl Acad Sci U S A* **96**: 15286-91.
- Visel A., Blow M.J., Li Z., Zhang T., Akiyama J.A., Holt A., Plajzer-Frick I., Shoukry M., Wright C., Chen F., Afzal V., Ren B., Rubin E.M., and Pennacchio L.A. (2009) ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**:854-8.
- Visel A., Thaller C., and Eichele G. (2004) GenePaint.org: an atlas of gene expression patterns in the mouse embryo. *Nucleic Acids Research* **32**:D552-D556.
- Vonsattel J.P., Myers R.H., Stevens T.J., Ferrante R.J., Bird E.D., and Richardson E.P., Jr. (1985) Neuropathological classification of Huntington's disease. *J. Neuropathol. Exp. Neurol.* **44**: 559-77.
- Voss A.K., Thomas T., and Gruss P. (1998) Efficiency assessment of the gene trap approach. *Dev. Dyn.* **212**:171-80.

- Wakeland E., Morel L., Achey K., Yui M., and Longmate J. (1997) Speed congenics: a classic technique in the fast lane (relatively speaking). *Immunol. Today* **18**:472-7.
- Walaas S.I., and Greengard P. (1984) DARPP-32, a dopamine- and adenosine 3':5'-monophosphate-regulated phosphoprotein enriched in dopamine-innervated brain regions. I. Regional and cellular distribution in the rat brain. *J. Neurosci.* **4**:84-98.
- Wallace D.M., Magnuson D.J., and Gray T.S. (1992) Organization of amygdaloid projections to brainstem dopaminergic, noradrenergic, and adrenergic cell groups in the rat. *Brain Res. Bull.* **28**:447-54.
- Wang L.P., Li F., Shen X., and Tsien J.Z. (2010) Conditional knockout of NMDA receptors in dopamine neurons prevents nicotine-conditioned place preference. *PLoS One* **5**:e8616.
- Wang W., Ni L., Yu Q., Xiong J., Liu H.C., and Rosenwaks Z. (2010) Expression of the *Lycat* gene in the mouse cardiovascular and female reproductive systems. *Dev. Dyn.* **239**:1827-37.
- Wang X., Devaiah S.P., Zhang W., and Welte R. (2006) Signaling functions of phosphatidic acid. *Prog. Lipid Res.* **45**:250-78.
- Wang Z., Kai L., Day M., Ronesi J., Yin H.H., Ding J., Tkatch T., Lovinger D.M., and Surmeier D.J. (2006) Dopaminergic control of corticostriatal long-term synaptic depression in medium spiny neurons is mediated by cholinergic interneurons. *Neuron* **50**:443-452.
- Wass M.N., Kelley L.A., and Sternberg M.J. (2010) 3DLigandSite: predicting ligand-binding sites using similar structures. *Nucleic Acids Res.* **38**:W469-73.
- Watkins L.H., Rogers R.D., Lawrence A.D., Sahakian B.J., Rosser A.E., and Robbins T.W. (2000) Impaired planning but intact decision making in early Huntington's disease: implications for specific fronto-striatal pathology. *Neuropsychologia* **38**: 1112-25.
- Wellcome Trust Case Control Consortium. (2007) Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**:661-78.
- West J., Tompkins C.K., Balantac N., Nudelman E., Meengs B., White T., Bursten S., Coleman J., Kumar A., Singer J.W., and Leung D.W. (1997) Cloning and expression of two human lysophosphatidic acid acyltransferase cDNAs that enhance cytokine-induced signalling responses in cells. *DNA Cell Biol.* **16**:691-701.
- Wexler N.S., Young A.B., Tanzi R.E., Travers H., Starosta-Rubinstein S., Penney J.B., Snodgrass S.R., Shoulson I., Gomez F., Ramos Arroyo M.A., et al. (1987) Homozygotes for Huntington's disease. *Nature* **326**:194-7.

- Wheeler V.C., Gutekunst C.A., Vrbanc V., Lebel L.A., Schilling G., Hersch S., Friedlander R.M., Gusella J.F., Vonsattel J.P., Borchelt D.R., and MacDonald M.E. (2002) Early phenotypes that presage late-onset neurodegenerative disease allow testing of modifiers in Hdh CAG knock-in mice. *Hum. Mol. Genet.* **11**:633-40.
- Wilson C.J., and Groves P.M. (1980) Fine structure and synaptic connections of the common spiny neuron of the rat neostriatum: a study employing intracellular inject of horseradish peroxidase. *J. Comp. Neurol.* **194**:599-615.
- Xenarios I., Rice D.W., Salwinski L., Baron M.K., Marcotte E.M., and Eisenberg D. (2000) DIP: the database of interacting proteins. *Nucleic Acids Res.* **28**:289-91
- Yaguchi T., Nagata T., and Nishizaki T. (2010) 1,2-dilinoleoyl-sn-glycero-3-phosphoethanolamine ameliorates age-related spatial memory deterioration by preventing neuronal cell death. *Behav. Brain. Funct.* **6**:52.
- Yamashita A., Nakanishi H., Suzuki H., Kamata R., Tanaka K., Waku K., and Sugiura T. (2007) Topology of acyltransferase motifs and substrate specificity and accessibility in 1-acyl-sn-glycero-3-phosphate acyltransferase 1. *Biochim. Biophys. Acta* **1771**:1202–1215.
- Ye G.M., Chen C., Huang S., Han D.D., Guo J.H., Wan B., and Yu L. (2005) Cloning and characterization a novel human 1-acyl-sn-glycerol-3-phosphate acyltransferase gene AGPAT7. *DNA Seq.* **16**:386-390.
- Yuki K., Shindou H., Hishikawa D., and Shimizu T. (2009) Characterization of mouse lysophosphatidic acid acyltransferase 3: an enzyme with dual functions in the testis. *J. Lipid Res.* **50**:860-9.
- Zala D., Benchoua A., Brouillet E., Perrin V., Gaillard M.C., Zurn A.D., Aebischer P., and Déglon N. (2005) Progressive and selective striatal degeneration in primary neuronal cultures using lentiviral vector coding for a mutant huntingtin fragment. *Neurobiol. Dis.* **20**:785-98.
- Zambrowicz B.P., Friedrich G.A., Buxton E.C., Lilleberg S.L., Person C., and Sands A.T. (1998) Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells. *Nature* **392**:608-11.
- Zeitlin S., Liu J.P., Chapman D.L., Papaioannou V.E., and Efstratiadis A. (1995) Increased apoptosis and early embryonic lethality in mice nullizygous for the Huntington's disease gene homologue. *Nat. Genet.* **11**:155-63.
- Zuccato C., Valenza M., and Cattaneo E. (2010) Molecular mechanisms and potential therapeutical targets in Huntington's disease. *Physiol. Rev.* **90**:905-981.

## ABSTRACT

This work concerns the functional study of two genes preferentially expressed in two brain regions affected by neurodegenerative diseases: Capucin, a marker of the striatum, a structure that degenerates in Huntington's disease and Agpat4, a marker of the ventral tegmental area and the *substantia nigra pars compacta*, whose dopaminergic neurons are selectively affected in Parkinson's disease. Mouse lines deficient for Capucin and Agpat4 have been generated in the laboratory and during my PhD thesis I carried out their characterization.

As the striatal gene expression of Capucin is significantly reduced in mouse models of Huntington's disease, we wished to evaluate its possible role in the pathogenesis of this disease. In a collaborative work, we examined the effect of the knockout and overexpression of the Capucin gene on the vulnerability of striatal neurons to a mutant Huntingtin fragment in a mouse model of Huntington's disease. The data show that Capucin has no significant effect on the toxicity of the mutant Huntingtin fragment in the considered model.

The Agpat4 protein has sequence homologies with acyltransferases involved in the metabolism of phosphoglycerides. I conducted expression studies using different molecular biology techniques, which showed that the Agpat4 gene is expressed in most catecholaminergic tissues. To determine the endogenous activity of Agpat4 and its physiological role in the tissues where it is expressed, I compared the metabolomes of Agpat4-deficient and wild-type mice tissues by liquid chromatography coupled with mass spectrometry. My results indicate that Agpat4 deficiency alters not only the metabolism of different lipid classes, in particular lysophosphatidylethanolamines, but also the metabolism of catecholamines.

## RESUME

Ce travail porte sur l'étude fonctionnelle de deux gènes préférentiellement exprimés dans deux régions du cerveau touchées par des pathologies neurodégénératives : Capucine, un marqueur du striatum, structure qui dégénère au cours de la maladie de Huntington et Agpat4, un marqueur de l'aire tegmentaire ventrale et de la substance noire compacte, dont les neurones dopaminergiques sont sélectivement atteints lors de la maladie de Parkinson. Des lignées de souris invalidées pour ces gènes ont été générées au laboratoire et au cours de ma thèse j'ai procédé à leur caractérisation.

L'expression striatale du gène de la Capucine étant significativement diminuée dans des modèles murins de la maladie de Huntington, nous avons souhaité évaluer son rôle éventuel dans la pathogenèse de cette maladie. Pour ce faire, nous avons examiné, dans le cadre d'une collaboration, l'effet du knock-out et de la surexpression du gène de la Capucine sur la vulnérabilité des neurones striataux à un fragment de la Huntingtine mutée dans un modèle murin de la maladie de Huntington. Les données montrent que la Capucine n'a pas d'effet significatif sur la toxicité du fragment de la Huntingtine mutée dans le modèle étudié.

La protéine Agpat4 présente des homologies de séquence avec des acyltransférases impliquées dans le métabolisme des phosphoglycérides. J'ai réalisé des études d'expression par différentes techniques de biologie moléculaire qui montrent que le gène d'Agpat4 est exprimé dans la plupart des tissus catécholaminergiques. Pour déterminer l'activité endogène d'Agpat4 et son rôle physiologique dans les tissus où elle est exprimée, j'ai comparé le métabolome de tissus de souris invalidées pour le gène d'Agpat4 et sauvages par chromatographie en phase liquide couplée à la spectrométrie de masse. Mes résultats indiquent que l'invalidation du gène d'Agpat4 perturbe le métabolisme non seulement de différentes classes de lipides, notamment les lysophosphatidyléthanolamines, mais aussi celui des catécholamines.